Progress and Challenges in Developing Tiered Quality Rating and Improvement Systems (TQRIS) in the Round 1 Race to the Top-Early Learning Challenge (RTT-ELC) States

November 2017

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education

This page has been left blank for double-sided copying.

Progress and Challenges in Developing Tiered Quality Rating and Improvement Systems (TQRIS) in the Round 1 Race to the Top-Early Learning Challenge (RTT-ELC) States

November 2017

Gretchen Kirby
Pia Caronongan
Andrea Mraz Esposito
Lauren Murphy
Megan Shoji
Patricia Del Grosso
Wamaitha Kiambuthi
Melissa Clark
Lisa Dragoset
Mathematica Policy Research

Tracy Rimdzius
Project Officer
Institute of Education Sciences

NCEE 2018-4003
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education
Betsy DeVos
Secretary

Institute of Education Sciences
Thomas W. Brock
Commissioner of the National Center for Education Research
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance
Ricky Takai
Acting Commissioner

November 2017

The report was prepared for the Institute of Education Sciences under Contract No. ED-IES-10-C-0077. The project officer is Tracy Rimdzius in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the reports.

Kirby, Gretchen, Pia Caronongan, Andrea Mraz Esposito, Lauren Murphy, Megan Shoji, Patricia Del Grosso, Wamaitha Kiambuthi, Melissa Clark and Lisa Dragoset. (2017). *Progress and Challenges in Developing Tiered Quality Rating and Improvement Systems (TQRIS) in the Round 1 Race to the Top-Early Learning Challenge (RTT-ELC) States* (NCEE 2018-4003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

This report is available on the IES website at http://ies.ed.gov/ncee.

CONTENTS

TABLES

FIGURES

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

Evidence about the benefits of high quality care for young children, particularly low-income children, has led to a strong commitment at the federal and state levels to improve the quality of early care and education. The Race to the Top—Early Learning Challenge (RTT-ELC) grants program is a joint effort of the U.S. Departments of Education (ED) and Health and Human Service (HHS) to strengthen the quality of early learning and development programs.

ED and HHS awarded RTT-ELC grants through three rounds of competition. Round 1 awards were made in December 2011 to nine states—California, Delaware, Maryland, Massachusetts, Minnesota, North Carolina, Ohio, Rhode Island, and Washington.

A central aspect of the RTT-ELC program is the development and implementation of tiered quality rating and improvement systems (TQRIS) as part of reforms to increase access to high quality, accountable programs, particularly for low-income and disadvantaged children. State RTT-ELC grant recipients were expected to make progress on five objectives related to TQRIS: (1) developing and adopting a common, statewide TQRIS; (2) promoting participation in the TQRIS; (3) rating and monitoring early learning and development programs, (4) promoting access to high quality programs for children with high needs by increasing the number of programs in the top levels of the TQRIS and increasing the number and percentage of children with high needs who are enrolled in programs that are in the top levels; and, (5) validating the effectiveness of the TQRIS.

The Department of Education's Institute of Education Sciences (IES) contracted with Mathematica Policy Research to conduct a descriptive study of TQRIS in the nine Round 1 RTT-ELC states four years after the receipt of grant funds. This report addresses two overarching questions:

1. How are TQRIS structured and implemented in the nine Round 1 RTT-ELC states?
2. How are TQRIS ratings defined, collected, and generated in the nine Round 1 RTT-ELC states?

The usefulness of the TQRIS ratings depends on their ability to measure program quality reliably and to construct a valid overall rating that distinguishes programs at different levels of quality (Cizek 2007, as cited in Zellman and Fiene 2012). Understanding the answers to the two questions above is important because TQRIS structure, implementation, and ratings procedures in the nine Round 1 states may have important implications for states' ability to reliably measure program quality and discern meaningful differences in quality across programs.

Data collection methods

To develop a comprehensive picture of TQRIS in the nine Round 1 RTT-ELC states, the study team used a master data collection protocol to assemble information incrementally across multiple data sources. The team first reviewed state RTT-ELC applications to gather foundational information. The team then reviewed publicly available and state-provided documents (such as program guides, TQRIS standards, and descriptions of TQRIS data systems) to update and expand upon the information collected from the applications. Last, the team

conducted interviews from January through April 2015 with administrators in the nine states who held in-depth knowledge about the TQRIS. The interviews filled remaining information gaps, confirmed important elements, and clarified any contradictions the team encountered in written materials.

Key findings

We summarize the findings from this descriptive study as they connect to the five RTT-ELC objectives for states listed above. This report focuses on the first three objectives. We first present a key finding that cuts across all three objectives, and then present key findings separately for each objective. Finally, although the fourth and fifth objectives were not the focus of this report, we mention below the other studies being conducted to shed light on whether states achieved those objectives, along with our findings on key features of states' TQRIS that may affect their ability to achieve those objectives.

There is tremendous variation across states in TQRIS structure, features, and processes; each state's TQRIS is uniquely designed and implemented. This key finding cuts across all three state objectives that were the focus of this report. States differ in terms of the timing of implementation, the policies used to promote participation, the amount and type of data available about programs and the children enrolled in them, the methods used to classify programs by type, the rating structure, the number of TQRIS components that contribute to the final rating, the way components are measured and defined, and how components are combined to arrive at the final rating.

Objective 1: Developing and adopting a common, statewide TQRIS

The nine Round 1 states made progress in developing and adopting a statewide TQRIS since the RTT-ELC grants were awarded. Four of the nine states had a statewide TQRIS in place in 2010, and that number increased to eight by 2015.

Objective 2: Promoting participation in the TQRIS

- Six of the nine Round 1 RTT-ELC states made TQRIS participation mandatory for certain programs (such as public prekindergarten programs or Head Start programs).
- The Round 1 RTT-ELC states promoted TQRIS participation and continuous quality improvement through incentives such as alternative pathways into high rating levels and financial incentives tied to higher ratings.

Objective 3: Rating and monitoring programs

- There were few commonalities across the nine states in the number of TQRIS components, the way they were defined, and how they were used to rate programs, even for the highest rating level.
- The nine Round 1 RTT-ELC states used different methods for calculating ratings, some of which changed over time. Three states require programs to meet all standards at a given level to achieve a particular rating. Others allow programs to choose, at least in part, which standards to pursue as they worked toward a particular rating. Four of the nine states changed their methods after receiving the RTT-ELC grant.

- TQRIS structure and policies may affect the distribution of programs across TQRIS rating levels. The maturity of the TQRIS and the decisions a state has made about its rating structure, policies, and incentives could affect the number of programs at each rating level.
- The nine Round 1 RTT-ELC states typically used various sources of evidence to collect information on TQRIS components, which could have potentially increased the reliability of the information. The nine states rarely relied exclusively on self-reported information from programs to construct ratings and typically verified self-reports with information from another source, such as on-site classroom observations.
- Six of the nine states required at least a bachelor's degree for assessors (staff who conducted classroom observations) and seven states required at least a bachelor's degree for raters (staff who verified other TQRIS components).
- Seven states required both assessors and raters to have professional experience in education.
- The Round 1 RTT-ELC states understood the importance of training and ensuring the reliability of staff who collected information for the TQRIS ratings. To prepare assessors, states used the training and reliability standards of the observation measure developers. States had to develop their own processes and standards to train and ensure reliability among raters.

Objective 4: Promoting access to high quality programs for children with high needs

Additional data are needed to fully understand the distribution of programs across the rating levels, increases in the number of programs in the top rating levels of TQRIS and the number of children with high needs enrolled in top-rated programs. As a follow-up to this descriptive study, the study team is working with administrators in the nine states to collect data on TQRIS participation and ratings from 2012 through 2016 to examine the extent to which participating programs moved from lower to higher rating levels.

Objective 5: Validating the effectiveness of the TQRIS

The nine Round 1 RTT-ELC states have all commissioned independent studies to validate the effectiveness of their TQRIS; these studies are expected to be completed in 2017. However, through this descriptive study of the nine states, we identified two key features of states' TQRIS that may affect the reliability and validity of their ratings:

- States faced trade-offs between promoting participation and quality improvement and producing valid, reliable ratings. These trade-offs are reflected in the structure of TQRIS in ways that can make it difficult to assess how accurately the ratings reflect program quality. For example, alternative pathways are intended to ease access and reduce the burden of participation among programs that have already met similar standards to those of the TQRIS. However, the underlying assumption that alternative pathway programs are comparable in quality to TQRIS rated programs at the same level has not been empirically confirmed.
- The complex composition of TQRIS ratings may weaken the relationship between ratings and child outcomes. TQRIS ratings in the nine states are based on many components. For some of those components (such as child-to-staff ratios and staff qualifications), there is

some evidence of associations between the components and improvement in child outcomes (such as language and math achievement). Other components (such as family engagement and community involvement) may be less directly related to children's classroom experiences, which may weaken the relationship between TQRIS ratings and child outcomes. The state validation studies may help address this by identifying the TQRIS rating components associated with child outcomes and ways to simplify the construction of ratings.

Conclusions

The nine Round 1 states made progress in developing and adopting statewide TQRIS since the RTT-ELC grants were awarded. The states worked to promote TQRIS participation among programs (particularly those that serve low-income children), define standards, and improve data collection and verification—with the goal of building systems that signal and promote quality. Follow-up work being conducted by Mathematica will examine the extent to which programs participating in TQRIS moved from lower to higher rating levels from 2012 through 2016. The states' validation studies, expected to be completed in 2017, may help shed light on the ability of TQRIS ratings to convey meaningful differences in program quality.

I. BACKGROUND AND MOTIVATION FOR THE STUDY

Evidence about the benefits of high quality early care and education (ECE) for young children has led to a strong commitment to quality improvement initiatives at the federal and state levels. The Race to the Top—Early Learning Challenge (RTT-ELC) grants program is a joint effort of the U.S. Departments of Education (ED) and Health and Human Service (HHS) to strengthen the quality of early learning and development programs.

ED and HHS awarded RTT-ELC grants through three rounds of competition. In the first round, 35 states, the District of Columbia, and Puerto Rico applied for RTT-ELC grants. Round 1 awards were made in December 2011 to nine states—California, Delaware, Maryland, Massachusetts, Minnesota, North Carolina, Ohio, Rhode Island, and Washington. Across these states, a total of $520 million dollars was awarded, with grants ranging from $45 million to $70 million (U.S. Department of Education RTT-ELC Awards).[1]

A central aspect of the RTT-ELC program is to support the development and implementation of tiered quality rating and improvement systems (TQRIS) as part of reforms to increase access to high quality, accountable programs, particularly for low-income and disadvantaged children (Applications for New Awards; Race to the Top—Early Learning Challenge 2011). To improve the quality of early learning and development, TQRIS aim to establish rigorous standards to define quality, rate programs based on those standards, and publicize the ratings of individual programs (Zellman et al. 2011). The nine Round 1 RTT-ELC states were judged by ED to have articulated comprehensive plans for, or made significant past progress in, developing and adopting a TQRIS based on a statewide set of standards that include early learning and development standards, assessment systems, health promotion, family engagement, and workforce competencies.

The history of quality rating systems pre-dates RTT-ELC. The first quality rating systems appeared in Oklahoma and Colorado in the late 1990s, primarily as a means of rating community-based programs that served children who received subsidized care and to promote consumer education among parents in choosing the care that their children received (Mitchell, 2005). The early rating systems were generally tied with tiered reimbursement programs in which states or localities paid increasingly higher subsidy rates for care in settings with higher ratings, especially those that achieved national accreditation. As the effort to rate quality in early care and education took hold across a growing number of states in the early 2000s, so did an emphasis on continuous quality improvement—expanding the systems beyond quality rating to quality rating and improvement (Zellman et al. 2008). The number of states nationwide with TQRIS has increased from 9 in 2004 to 23 in 2010 (Tout et al. 2010), just before the launch of RTT-ELC. Four of the nine Round 1 RTT-ELC states were implementing statewide TQRIS prior to receipt of the RTT-ELC grant and another three states were implementing pilots. By 2015, 40 states had a TQRIS (The BUILD Initiative and Child Trends 2015).

[1] California received $53 million in 2011 and a supplementary grant of $22 million in 2013.

State RTT-ELC grant recipients were expected to make progress in five areas related to TQRIS: (1) developing and adopting a common, statewide TQRIS; (2) promoting participation in the TQRIS; (3) rating and monitoring early learning and development programs; (4) promoting access to high quality programs for children with high needs by increasing the number of programs in the top levels of the TQRIS and increasing the number and percentage of children with high needs who are enrolled in programs that are in the top levels; and (5) validating the effectiveness of the TQRIS.

The Department of Education's Institute of Education Sciences (IES) contracted with Mathematica Policy Research to conduct a descriptive study of TQRIS in the nine Round 1 RTT-ELC states four years after the receipt of grant funds. The study focused on TQRIS in the context of center- (or school-) based early learning and development programs serving preschool children age 3 to 5 years. The reason for this focus was that a component of the study explored the feasibility of using administrative data to examine the relationship between TQRIS ratings of these programs serving children in the year prior to kindergarten and child outcomes measured at kindergarten entry.[2]

A. Research questions

In this report, we present findings from a descriptive study of two overarching questions:

1. How are TQRIS structured and implemented in the nine Round 1 RTT-ELC states?
2. How are TQRIS ratings defined, collected, and generated in the nine Round 1 RTT-ELC states?

The usefulness of the TQRIS ratings depends on their ability to measure program quality reliably and to construct a valid overall rating that distinguishes programs at different levels of quality (Cizek 2007, as cited in Zellman and Fiene 2012). If TQRIS ratings do capture true differences in quality, then these differences might be reflected in the development of social and behavioral skills, or literacy, language and math skills for children who attend programs with lower versus higher ratings. Understanding the answers to the two research questions above is important because TQRIS structure, implementation, and ratings procedures in the nine Round 1 states may have important implications for states' ability to reliably measure program quality and discern differences in quality across programs that can ultimately result in differential outcomes for children's school readiness.

[2] The feasibility component found that a correlational analysis of the relationship between TQRIS ratings and kindergarten entry assessment scores could be feasible in three of the Round 1 RTT-ELC states but, due to data limitations in the coverage of children and programs, the analysis would have limited utility. Findings from the feasibility assessment are available upon request.

B. Data collection methods

To develop a comprehensive picture of TQRIS in the nine Round 1 RTT-ELC states, the study team used a master data collection protocol to assemble information incrementally across multiple data sources.[3] The team first reviewed state RTT-ELC applications to gather foundational information. The team then reviewed publicly available and state-provided documents (such as program guides, TQRIS standards, and descriptions of TQRIS data systems) to update and expand upon the information collected from the applications. Last, the team conducted interviews from January through April 2015 with administrators in the nine states who held in-depth knowledge about the TQRIS. The interviews filled remaining information gaps, confirmed important elements, and clarified any contradictions the team encountered in written materials.

C. Roadmap to the report

This report focuses on the progress made and challenges faced by the nine states in achieving the first three objectives for which they were held accountable for use of their RTT-ELC funds (listed above). We discuss discrete findings related to these objectives in relevant chapters of the report: we discuss TQRIS implementation and participation in Chapter II, TQRIS structure in Chapter III, and the process to generate TQRIS ratings in Chapter IV. In Chapter V, we review the refinements Round 1 states made to their TQRIS since receipt of the RTT-ELC grant. Although the fourth and fifth state objectives listed above were not the focus of this report, in Chapter VI, we describe the other studies that are being conducted to shed light on whether states achieved these objectives, along with our findings on key features of states' TQRIS that may affect their ability to achieve those objectives. Appendix A contains profiles for each of the nine states.[4]

Notable from this study, as with previous TQRIS cross-state work (Kirby et al. 2015; Tout et al. 2010; Zellman et al. 2008), is the tremendous variation in the structure, features, and processes across the TQRIS; each state TQRIS is uniquely designed and implemented. Nonetheless, the TQRIS decisions and implementation experiences of the nine Round 1 RTT-ELC states can help guide federal and state efforts to use TQRIS to improve the quality of early learning and development programs. In particular, the findings presented in this report may illuminate potential areas of strength and weakness in the ability of states' TQRIS to identify differences in program quality that may be meaningful for children's outcomes.

[3] Mathematica Policy Research (2017) contains the master protocol used by the study team to collect and organize the information from documents and interviews.

[4] Although the main body of this report focuses on TQRIS in the context of center- (or school-) based early learning and development programs serving preschool children age 3 to 5 years, some of the state profiles in Appendix A present information (such as the distribution of programs across rating levels) for a wider range of programs, because the information provided by those states was not disaggregated by program type.

This page has been left blank for double-sided copying.

II. TQRIS IMPLEMENTATION AND PARTICIPATION

In this chapter, we discuss the timing of TQRIS implementation—that is, when the nine Round 1 RTT-ELC states launched their TQRIS and began operating the systems, the eligibility of center-based programs, and the incentives to promote TQRIS participation and advancement. These aspects of TQRIS provide important context for understanding the ability of ratings to signal meaningful differences in quality.

A. Timing of TQRIS implementation in the Round 1 RTT-ELC states

The nine Round 1 states made progress in developing and adopting common, statewide TQRIS since the RTT-ELC grants were awarded. In 2010, just four of the nine states had a statewide TQRIS, and three others had pilots under way (Table II.1). By early 2015, all but one of the Round 1 RTT-ELC states were implementing TQRIS statewide. In the one remaining state (California), TQRIS was locally administered but had not been adopted in all areas (and we therefore classify it as being in a pilot phase). The states varied in how long they had been using TQRIS. Across the 9 states, TQRIS had been implemented in some form for an average of seven years. One state began using TQRIS in 1999; the other states did not begin implementing TQRIS statewide until 2006 or later. The difference in start dates reflects a broader trend in TQRIS implementation across the United States. Since 2005, TQRIS has gained momentum nationally as an approach to quality improvement in early childhood education (Zellman and Perlman 2008).

Table II.1. Timing of TQRIS implementation in Round 1 RTT-ELC states

State	Name of TQRIS	Years of pilot implementation[a]	Year of statewide implementation
California	Race to the Top - Early Learning Challenge Grant Tiered Quality Rating and Improvement System	16 counties began in 2012	n.a.
Delaware	Delaware Stars for Early Success	2007	2008
Maryland	Maryland EXCELS	2012	2013
Massachusetts	Massachusetts Quality Rating and Improvement System	2010-2011	2011
Minnesota	Parent Aware	2007-2011	2015
North Carolina	Star Rated License System	n.a.	1999
Ohio	Step Up to Quality Program	2004-2005	2006
Rhode Island	BrightStars	2007-2008	2009
Washington	Early Achievers	2008-2011	2012

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

[a] County-administered programs that are not statewide are included as pilots.

n.a. = not applicable.

Eight of the Round 1 RTT-ELC states piloted TQRIS before expanding statewide. Pilots provided states with an opportunity to implement the standards, the rating determination process, and quality improvement approaches with a subset of early childhood education programs or in limited geographic areas. Pilots lasted one to 5 years, with an average length of 2.5 years.

B. TQRIS eligibility for center-based programs

The selection criteria for RTT-ELC grants encouraged states to adopt policies and practices that promote TQRIS participation among all publicly funded early learning and development programs (Applications for New Awards; Race to the Top—Early Learning Challenge 2011). Eligibility is the word that states use to indicate which programs are allowed to participate in the TQRIS. Some eligible programs may be required to participate—meaning that participation is mandatory for these programs—while other eligible programs have the option to participate or not—meaning that participation is voluntary for these programs.

In all nine Round 1 states, most center-based programs serving 3- to 5-year-old children were eligible to participate in TQRIS, including public prekindergarten, Head Start, and licensed private center-based child care programs (Table II.2). In four states, license-exempt private center-based child care programs were also eligible to participate.[5] Therefore, the only center-based programs not eligible to participate in TQRIS were those that were license-exempt in five states (Delaware, Minnesota, Ohio, Rhode Island, and Washington).

Six of the nine Round 1 RTT-ELC states made TQRIS participation mandatory for certain programs. In three states, participation in TQRIS was voluntary for all eligible programs. In the other six states, TQRIS participation was mandatory for some but not all eligible programs. Four states required TQRIS participation of public prekindergarten programs, and three of these four also required participation of Head Start programs. In one state (North Carolina), all licensed programs were required to enroll in TQRIS at level 1 and then could choose to apply for a higher rating level. In this state, along with three others, center-based programs receiving subsidies from the Child Care and Development Fund to provide care for low-income children had to participate in TQRIS. Participation of license-exempt centers, in the four states in which they were eligible, was always voluntary.

Overall, the states increased their participation requirements under the RTT-ELC grants. Of the six states out of the nine for which information is available, only North Carolina reported any participation requirements in the TQRIS in 2010, before the grants were awarded (Tout et al. 2010).

[5] License-exempt programs are legally operating but exempt from state licensing requirements. Depending on the state, they can include programs run by or affiliated with religious organizations or programs operated by public schools.

Table II.2. TQRIS eligibility and participation requirements for programs serving 3- to 5-year-old children in Round 1 RTT-ELC states, by program type

Program type	Number of states in which programs are eligible to participate in TQRIS	Number of states in which participation is:		
		Voluntary for all programs	Mandatory for some programs	Mandatory for all programs
All center-based programs	9	3	6	0
Public pre-kindergarten	9	5	0	4
Head Start	9	6	1[a]	2
Licensed center-based programs	9	4	4[b]	1[c]
License-exempt centers	4	4	0	0

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

[a] The state-funded Head Start equivalent program in Delaware, called the Early Childhood Assistance Program (ECAP), made participation mandatory.

[b] TQRIS participation is mandatory for licensed center-based programs receiving funding from the Child Care and Development Fund in all four states.

[c] TQRIS participation at level 1 is mandatory for all licensed center-based programs in North Carolina and voluntary for higher rating levels.

Six states reported outreach efforts aimed at further expanding TQRIS participation among center-based programs for which participation was not mandatory. Two of the six states targeted programs that received CCDF for increased TQRIS participation. Two other states among the six reported efforts to increase participation among part-day preschool programs and public prekindergarten programs. In a fifth state, respondents reported building participation among community-based child care programs in general. In the sixth state, California (the only state not already operating statewide), respondents discussed intentions to expand TQRIS to additional counties.

States faced challenges in gaining the participation of certain programs in TQRIS, particularly programs that had to meet other standards. For example, two states reported challenges in gaining the participation of Head Start programs, and two states described difficulties in expanding participation among nationally accredited licensed center-based programs (most often with an accreditation from the National Association for the Education of Young Children [NAEYC], a national accreditation system for early childhood education programs). Head Start programs had to meet Head Start Program Performance Standards to maintain federal funding, and some licensed center-based programs voluntarily pursued national accreditation for providing a high-quality early learning and development program. In these cases, the programs opted not to also pursue TQRIS ratings. In contrast, in one state, administrators found that some licensed center-based programs were opting to pursue high-level TQRIS ratings in lieu of pursuing renewal of national accreditation.

Administrators in four states also noted challenges to encouraging TQRIS participation among school-based public prekindergarten programs. In three of the four states, administrators described challenges in using the Early Childhood Environment Rating Scale (ECERS; Harms, Clifford, and Cryer 2005) as part of the rating determination for these programs because the layout of school buildings can result in lower ECERS scores. One of the three states also noted

that school-based public prekindergarten programs used a data system other than that associated with TQRIS and followed a different professional development model for their staff, making TQRIS participation time-consuming and potentially poorly aligned with current practice. One state (North Carolina) made attainment of a TQRIS level 4 rating an eligibility requirement for receipt of funding through the state-funded prekindergarten program.

C. Incentives to promote TQRIS participation and advancement

The Round 1 RTT-ELC states used supports and incentives to promote TQRIS participation and to foster continuous quality improvement. These included alternative pathways into high rating levels and financial incentives to encourage programs to earn higher ratings.

1. Alternative pathways

To encourage participation, six states used alternative pathways into high rating levels for certain program types (Figure II.1). Alternative pathways allowed programs to meet standards automatically or receive a rating without going through the full data collection and verification process for TQRIS standards. Two of the six states added automatic upper-level ratings for public prekindergarten and Head Start programs since receiving the RTT-ELC grant.

Figure II.1. Use of alternative pathways to TQRIS ratings in Round 1 RTT-ELC states

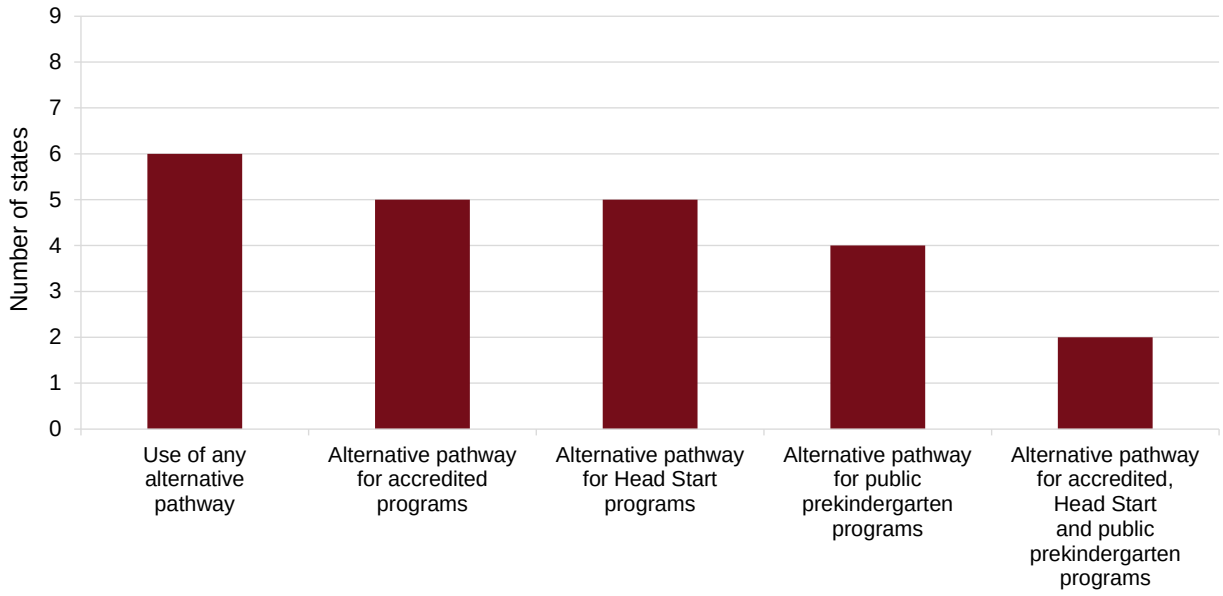


Exhibit reads: Six states used an alternative pathway to determine TQRIS ratings.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

In all six states, alternative pathways were available only to programs meeting specific sets of external standards that were often similar to TQRIS standards. For example, five of the six states applied alternative pathways to programs with an existing accreditation. Five of the six offered alternative pathways to Head Start programs and four extended the alternative pathways to public prekindergarten programs. Two states (Delaware and Minnesota) offered alternative pathways to each type of program—accredited, Head Start programs, and public prekindergarten. The use of alternative pathways aimed to reduce duplication of effort to demonstrate similar standards and to overcome the associated barriers to participation.

In three of the six states, the alternative pathway automatically assigned a given rating level to eligible programs. For example, in Delaware, Head Start programs automatically received a level 4 rating on a scale of 1 through 5. In the other three states, eligible programs automatically received credit for meeting some standards but still had to demonstrate adherence to additional TQRIS-specific standards. For example, in Rhode Island, Head Start and accredited programs received automatic credit as level 5 on 3 of the state's 10 standards but independently had to demonstrate the other 7 standards. Programs entering TQRIS through an alternative pathway typically entered at the mid- to high-level rating, but not always the highest. The programs could apply to achieve a higher level by demonstrating the necessary TQRIS standards that extended beyond those specified in the alternative pathway. The additional standards typically involved classroom observations.

2. Financial incentives

Every Round 1 RTT-ELC state used some form of a financial incentive or combination of incentives to promote TQRIS participation and advancement up the rating levels. Seven states offered grants to programs participating in TQRIS to support quality improvement (Table II.3). States made the grants available to programs at all levels in six of the seven states and to programs at each level above the lowest level in the remaining state. TQRIS in six states provided monetary awards or bonuses to programs that achieved a specific rating level: three states provided such awards to programs for reaching each rating level, two provided bonuses to programs for achieving only upper-level ratings, and one provided the bonus to programs for achieving each level but the highest. Two states offered monetary awards or bonuses for staff from TQRIS-participating programs; in one state, the bonuses rewarded staff for achieving specified credentials. Three states also offered scholarships for staff; in two states (Rhode Island and Washington), the scholarships were exclusively for staff from TQRIS-participating programs and in Washington the scholarships were only available to staff in programs with a level 3 rating or higher. The third state (Delaware) offered staff from TQRIS participating programs priority in funding decisions. Seven states offered increased rates per child for subsidized care through CCDF as the TQRIS rating level increased (referred to as tiered reimbursement).

Table II.3. Incentives offered to programs participating in TQRIS in Round 1 RTT-ELC states, by incentive type

Incentive type	Number of states in which incentive was available	Level at which incentive is available (number of states)	Eligibility[a]
Quality improvement grants	7	All levels (six states) All above level 1 (one state)	Restricted to TQRIS-participating programs
Bonus award for achieving a rating level	6	All levels (three states) Upper levels (two states) All but highest level (one state)[b]	Restricted to TQRIS-participating programs
Monetary awards or bonuses for staff	2	All levels (two states)	Restricted to TQRIS-participating programs
Scholarships	3	All levels (two states) Upper levels (one state)	Restricted to staff in TQRIS-participating programs in two states; staff from TQRIS-participating programs received priority in one state
Tiered reimbursement	7	All levels (two states) Upper levels (five states)	Restricted to TQRIS-participating programs[c]

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

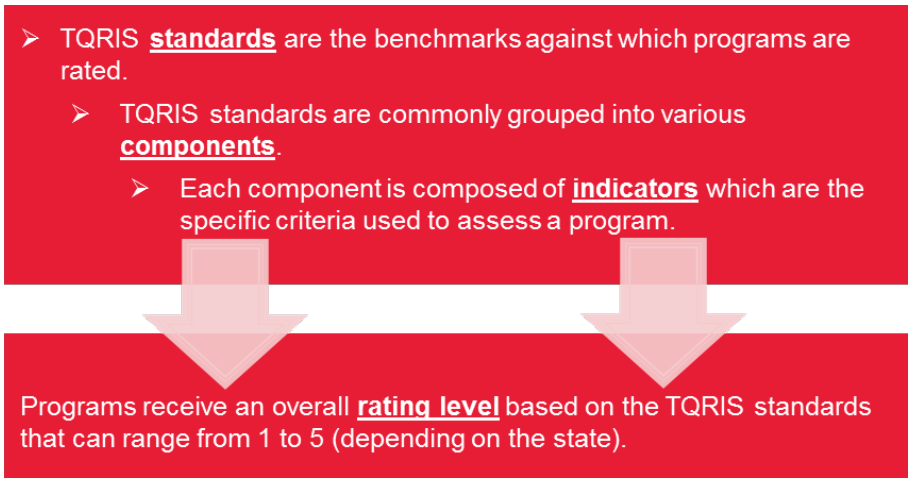
[a] Eligibility applies to all states in which the incentive was offered.

[b] Minnesota provided post-rating support of up to $1,000 to programs after they achieved levels 1, 2, or 3.

[c] Massachusetts provided tiered reimbursement to programs that were at level 2 or above and served infants and toddlers.

III. TQRIS STRUCTURE

The definition of quality, as determined by TQRIS standards and the structure of the rating system, potentially influences the ability of TQRIS ratings to make meaningful distinctions in program quality. In this chapter, we first describe how the nine Round 1 RTT-ELC states defined quality through TQRIS standards and determined rating levels. We then present an illustration of the distribution of programs across the rating levels. Lastly, we examine whether the states' TQRIS included components that research has shown to be associated with child outcomes (such as language and math achievement).



A. Defining quality

TQRIS define quality through a set of standards that the systems measure to signal to programs and parents the elements of high quality early child care and education. The usefulness of TQRIS standards depends on how well they can measure true differences in program quality. TQRIS standards are commonly grouped into categories or components (QRIS National Learning Network 2014). The TQRIS components vary across states in response to different emphases on particular goals for children, families, programs, and early childhood education systems (Kirby et al. 2015; Zaslow and Tout 2014).

No two states defined TQRIS components or included them in the rating structure in the same way. There were few commonalities across the nine states in the number of TQRIS components, the way they were defined, and how they were used to rate programs, even for the highest rating level.

We used an existing framework of 13 quality components from the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010) to categorize the TQRIS components across the nine states.[6] As shown in Figure III.1, all nine states included 3 of the components—licensing, quality of the environment, and workforce qualifications and credentials—and all but one state included the child assessment, use of developmentally-appropriate curriculum, and

[6] Similar categories are used in the online QRIS Compendium maintained by the BUILD Initiative (The BUILD Initiative and Child Trends, 2016).

family engagement components. Five states included 10 or more of the 13 TQRIS components (not shown in figure). Five states included the health promotion and safety component, consistent with earlier research showing that states often rely on licensing standards to ensure health and safety and do not impose additional requirements for health and safety in TQRIS (Kirby et al. 2015). The four states that included the cultural and linguistic diversity component also included two other less common TQRIS components—community involvement and provisions for special needs.

Not every component underwent evaluation at each TQRIS rating level. For example, licensing entered the rating process as a prerequisite or a requirement at the lowest level while accreditation typically entered as a requirement at the highest levels (level 4 or 5). The other TQRIS components demonstrated considerable variation in the level at which requirements entered the rating process (Figure III.1). For example, even though all nine TQRIS included environment as a component, requirements for environment entered the rating at level 1 in three states, level 2 in three states, and level 3 in three states. Requirements at lower levels tended to be less stringent than those at higher levels. For example, states with requirements for environment at the lower rating levels specified environment as a self-assessment using a classroom observation measure and only required an independent, scored assessment at higher rating levels.

Figure III.1. Number of states with requirements for each component and level at which requirements entered the rating process for Round 1 RTT-ELC states

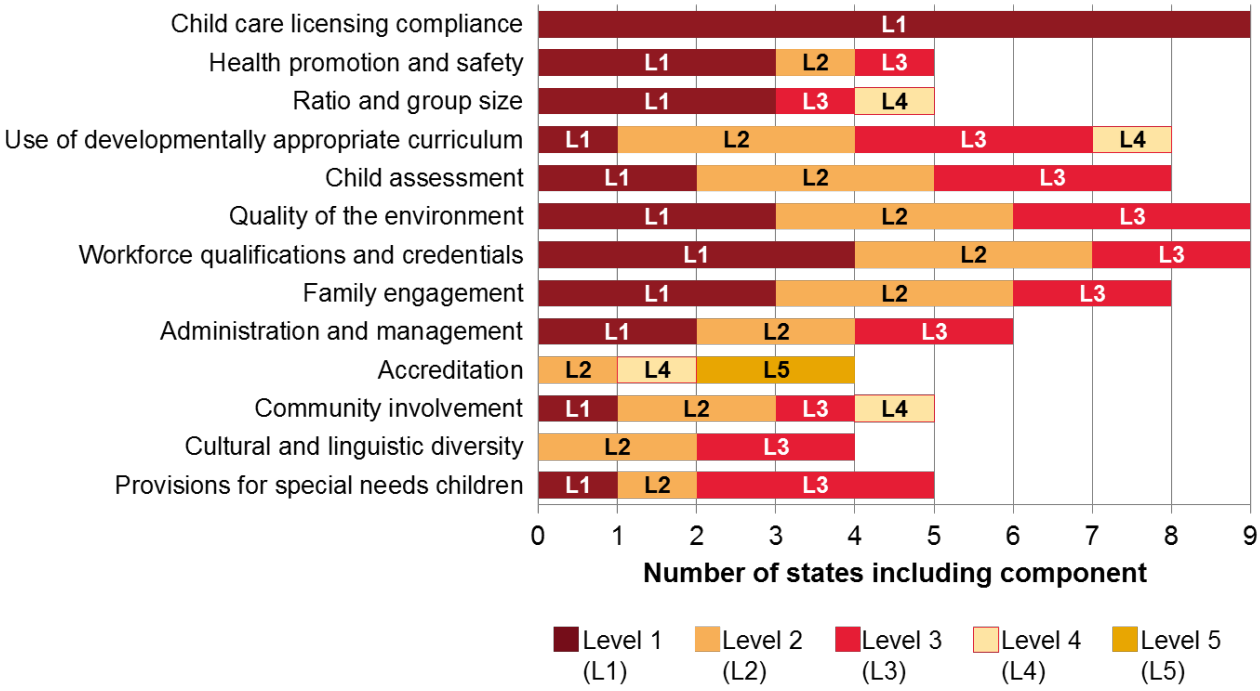


Exhibit reads: All nine states included requirements for the child care licensing compliance component at Level 1; three states included requirements for the health promotion and safety component at Level 1, one state included health promotion and safety requirements at Level 2, and one state included these requirements at Level 3.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

We found wide variation across the nine states in the components assessed for the lowest and highest TQRIS rating levels. At the lowest rating level, the number of components ranged from 1 to 7 across the states. In three states, licensing compliance was the only required component to achieve the lowest rating level. At the highest rating level, the number of assessed components ranged from 6 to 12. For some components, states set requirements at each rating level and increased the rigor at each higher level. For example, Ohio's TQRIS included a requirement for the curriculum component at each level. To achieve a level 1 rating, programs had to begin the process of identifying a curriculum that was aligned with state early learning standards; to achieve a level 3 rating, programs had to be implementing that curriculum; to earn points towards a level 4 or 5 rating, programs had to be implementing activities aligned with that curriculum across all developmental domains. All states increased the number of required components as the TQRIS rating level increased. For example, Maryland had requirements in place for 4 components at level 1 but 12 components at level 5.

Each TQRIS component is composed of indicators, which are the specific requirements a program must meet. Required indicators for the highest rating level were more similar for some components than for others. For example, for the curriculum component, eight states required programs to implement a curriculum aligned with state early learning standards in order to achieve the highest rating. In contrast, requirements for administration and management were not consistent at the highest level across states. Six states required paid planning time for staff, three required inspection of an annual budget, and three required evidence of procedures that supported continuous quality improvement. None of the six states with requirements for the administration and management component required all three of these elements at the highest level.

B. Defining rating levels

States make decisions about the structure of TQRIS ratings and the distinctions between levels based on expert opinion, stakeholder input, and programs' ease of entry and progression up the rating levels (Kirby et al. 2015). States use different methods to combine the indicator data to assign component-level ratings and determine the overall rating level a program receives (The BUILD Initiative and Child Trends, 2016; QRIS National Learning Network 2014). TQRIS typically use one of three rating structures:

1. A ***building-block structure*** requires a program to meet all the quality indicators for every component at a given level to receive the rating for that level.
2. A ***points structure*** allows a program to earn points for indicators in each component of the TQRIS standards, with the total number of points determining the program's rating level.
3. A ***hybrid structure*** combines elements of the building-block and points structures by, for example, determining lower rating levels with a building-block structure and higher rating levels with a points structure.

These three rating structures each offer programs a different degree of flexibility in meeting TQRIS standards. Points or hybrid structures give programs more flexibility in demonstrating quality, which presumably encourages programs to earn higher ratings. In contrast, a building-block structure conveys the importance of specific requirements at each rating level and equalizes the definition of quality across programs.

The nine Round 1 RTT-ELC states used a variety of TQRIS rating structures. Three states used a building-block structure, two used a points structure, and four used a hybrid structure (Figure III.2). All four states with a hybrid structure used building blocks for the first one to three rating levels and points for the higher rating levels. Seven states used five rating levels; the other two states used four.

Figure III.2. Decisions about TQRIS structure, number of rating levels, and time frame for which ratings were valid in Round 1 RTT-ELC states

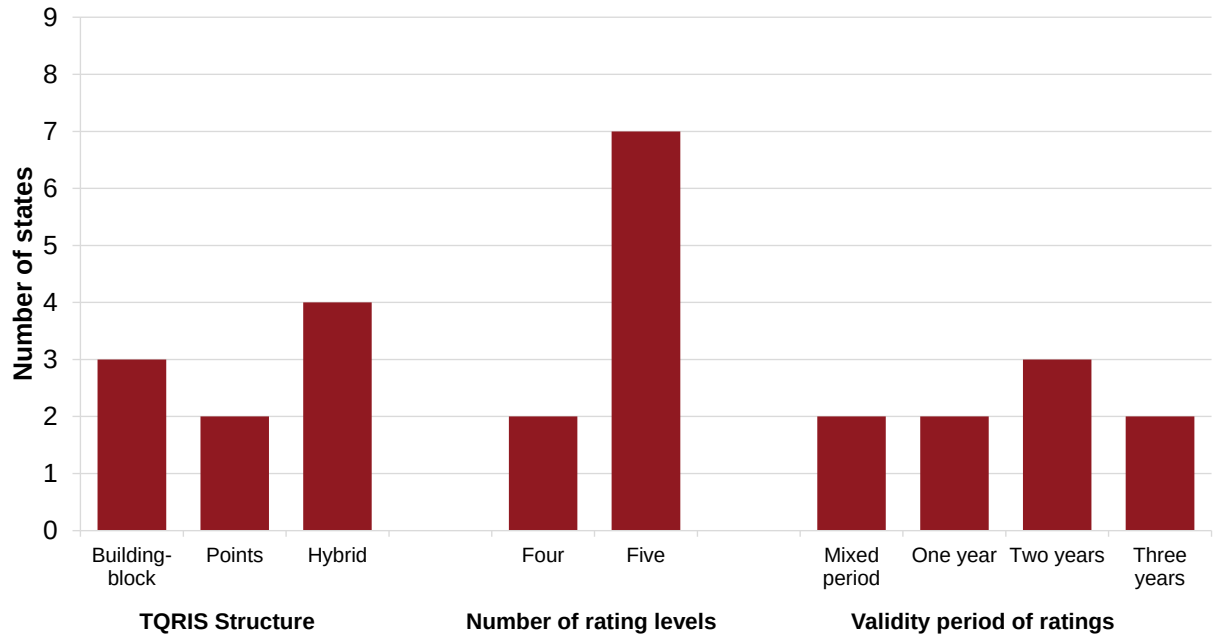


Exhibit reads: Three states used a building-block TQRIS structure, two states used a points structure, and 4 states used a hybrid structure.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

1. Entry into and advancement through the TQRIS rating levels

The policies that determined how programs entered and progressed up the rating levels varied by state. Four states required programs to enter at the first level while the other five states allowed programs to apply to enter at any level. For example, a program seeking to participate in TQRIS for the first time could apply for a rating level 3. After initial entry, all nine states allowed programs to receive any rating level for which they were determined qualified. For example, a program with a rating level 1 could increase to a rating level 3. Only in one state—Massachusetts—were all programs encouraged (but not required) to progress sequentially through each rating level.

Four of the nine Round 1 RTT-ELC states had policies about how quickly programs could advance from one rating level to the next; at a minimum, programs had to remain at a given level for three months to one year. In other states, the timing was determined by either each program's initiative in applying for a higher level or the frequency with which participating programs underwent re-evaluation, as set by the state. The information reported by state respondents suggested that the time required to progress between levels could increase as programs moved up

the levels; for example, it became increasingly difficult for programs to achieve minimum required scores on classroom observation measures (reported by two states) or to achieve staff education or credentialing requirements at the higher levels (reported by five states). Only one state (Delaware) imposed a limit on the maximum amount of time a program could remain at a particular level; programs had to advance from level 1 within one year and level 2 within two years or be removed from the system and encouraged to re-apply.

2. Validity period of ratings

Within each state, programs participating in TQRIS underwent re-evaluation within a specified period if they did not initiate an application for a new rating. The re-evaluation often led to a program's receipt of the same rating level but could result in a change in rating level (however, state respondents said that a decrease in the rating level was rare). In two states (Delaware and Ohio), ratings were valid for three years for the highest rating levels (levels 3 through 5 in Delaware and 4 and 5 in Ohio), but shorter timeframes were set for the lower levels. Ratings were valid for one year at level 1 and two years at level 2 in both states. Ratings for level 3 were also valid for two years in Ohio. Two states re-evaluated all rating levels at least annually, three at least every two years, and the remaining two states at least every three years (Figure III.2). However, several factors could contribute to more frequent system-initiated re-evaluations. For example, all nine states required a rating redetermination if a program had a licensing compliance issue. Other changes, such as a change in program director, location, configuration of classrooms, or the ages of children served, could also trigger a rating re-evaluation in each state.

C. Distribution of programs across TQRIS rating levels

Direct comparisons of programs at each TQRIS rating level across states are not possible because of the differences in state definitions of each rating level and their methods of counting the number of programs at each level. For example, a level 3 rated program in one state may not have a comparable level of quality to a level 3 rated program in another state because of differences in the definitions of the levels and the way the ratings are calculated.

Although rating levels cannot be compared across states, it is appropriate and informative to compare states with regard to the distribution of their programs across TQRIS rating levels. Figure III.3 illustrates the great deal of variation that existed in the nine states in the distribution of programs across rating levels. This variation is indicative of the differences among the states in several factors that likely affected how programs were distributed across rating levels, including differences in the participation patterns of certain programs, the availability of alternative pathways (discussed above), the number of rating levels, and policies about entry into and advancement through the rating levels (the latter two are discussed in Chapter III), the maturity of TQRIS, and the structure of the TQRIS standards. One limitation of Figure III.3 in making cross-state comparisons is that some states (specifically, Delaware, Massachusetts, Minnesota, North Carolina, Ohio, and Washington) provided data on the distribution of programs across rating levels for all programs combined, while other states (California, Maryland, and Rhode Island) provided data only for center-based programs. With this limitation in mind, Figure III.3 can still be used to gain a general sense of the wide variation across states with regard to how many programs fall into each rating level.

For example, Massachusetts and Minnesota had dramatically different distributions of programs across TQRIS levels with ratings skewed to the low end in Massachusetts and the high end in Minnesota; the vast majority of programs in Massachusetts are at Level 1 whereas the vast majority of programs in Minnesota are at Level 4. This could be because of key differences between these two state's TQRIS, such as the fact that Massachusetts had a building block system that encouraged programs to enter at level 1 and progress through each rating whereas Minnesota had a hybrid system that allowed programs to apply for any rating level at entry.

TQRIS structure and policies may affect the distribution of programs across TQRIS rating levels. The maturity of the TQRIS and the decisions a state has made about its structure and implementation could affect the number of programs at each rating level. These decisions include the rating structure as well as policies and incentives to encourage programs to participate and earn higher ratings. Such decisions and policies can introduce noise—that is, a program's rating may not just reflect quality but also the program's response to the system's structure and incentives.

Figure III.3. Distribution of rated programs across TQRIS rating levels in Round 1 RTT-ELC states

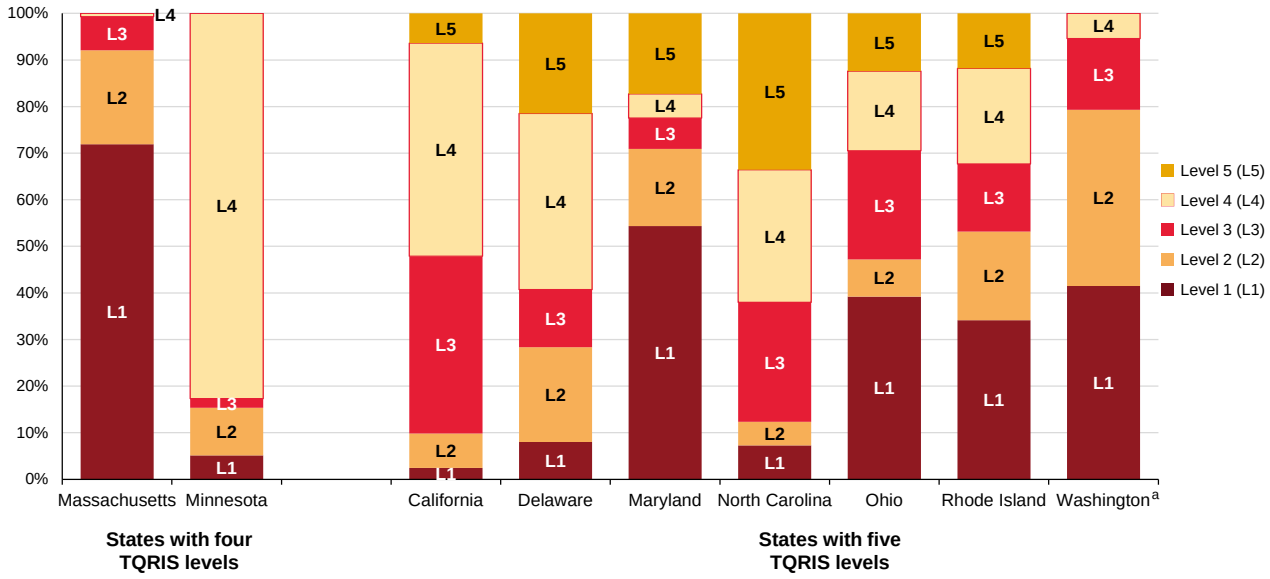


Exhibit reads: In Massachusetts 72 percent of programs were rated at Level 1, 20 percent were rated at Level 2, 7 percent were rated at Level 3, and 1 percent were rated at Level 4.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

Note: Distributions calculated by Mathematica based on the most recent information on participation by level as provided by each state or in the RTT-ELC 2014 Annual Performance Reports (APRs). Data for Delaware, Massachusetts, Minnesota, North Carolina, Ohio, and Washington are not restricted to center-based programs. States vary in their definition of TQRIS participation and in their methods for counting the number of programs at each level. Further information is provided in each state profile in Appendix A.

[a] Washington state has five TQRIS levels but less than one percent of programs reached the highest level.

D. Presence of TQRIS components associated with child outcomes in the research literature

Participation in high-quality ECE has been linked to positive outcomes for children, such as language and math achievement. Researchers have sought to identify which specific features of ECE produce these benefits. Features of high-quality ECE that have demonstrated associations with child outcomes, particularly in the domains of literacy and math, include child-to-staff ratios and group size, curriculum, staff qualifications, and environment (as measured by observed quality and teacher-child interactions) (Burchinal et al. 2010; Yoshikawa et al. 2013; Zaslow et al. 2010). It is important to note that most of the evidence pertaining to the relationship between specific features of ECE and child outcomes is based on correlational studies and that many of these studies suggest that the contribution of any one feature tends to be small (Caronongan et al. 2016). Nonetheless, given their prominence in the research base, we examined the extent to which these features were included as components in the Round 1 states' TQRIS. All nine states included staff qualifications and environment. These two components have been a staple in even the earliest systems developed over a decade ago (Zellman and Perlman 2008).

The indicators used to measure child-to-staff ratios and group size, curriculum, and environment were similar across the states that included them in the TQRIS (Table III.1). Indicators for the staff qualifications component differed and included five potential indicators: percentage of staff with a specific education level, credential, training, years of experience, or meeting state-defined career lattice requirements.[7]

Table III.1. Indicators to measure TQRIS components that the research literature has shown to be associated with child outcomes

Component	Number of states with component	Indicator to measure component	Number of states with indicator
Child-to-staff ratios and group size for 3- to 5-year-old children	5	Maximum ratios based on ages of children	5
		Maximum group size per classroom based on ages of children	3
Use of developmentally appropriate curriculum	8	Selection and use of curriculum aligned with state early learning standards	8
		Teacher training in curriculum	4
Quality of the environment	9	Independent observational assessment	9
Staff qualifications	9	Percentage of teaching staff with required education level	7
		Percentage of teaching staff with required credential	6
		Percentage of teaching staff with required training in specific topics	3
		Percentage of teaching staff with required years of experience	1
		Percentage of teaching staff meeting state career lattice requirements	2

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

[7] A career lattice defines staff qualification requirements that are used among practitioners across a state's early learning and development programs.

Three of the nine states placed additional emphasis on components associated with child outcomes by giving such components more weight in the calculation of the overall rating. For example, Delaware awarded 60 out of 100 total points for staff qualifications and environment. Washington awarded 80 out of 100 total points for the overall rating across curriculum, environment, and staff qualifications, with 55 percent of the overall rating based on an observational assessment of the environment. The greater weight assigned to these components makes a program's performance on these components more important relative to performance on other components.

Observational assessment of classrooms are a key indicator of the environment component in TQRIS. All nine Round 1 RTT-ELC states used classroom observation measures in their TQRIS to assess classrooms with preschool age children (3- to 5-year-olds). The measures were typically the Classroom Assessment Scoring System (CLASS; Pianta, LaParo, and Hamre 2008) or the Early Childhood Environment Rating Scale, Revised Edition (ECERS-R; Harms, Clifford, and Cryer 2005) (Figure III.4). Four states used both measures, but only for specific rating levels or program types. Rhode Island, for example, required the ECERS-R for any center-based program applying for level 3 or higher and then added the CLASS only for those applying for level 5. Maryland used the CLASS in public prekindergarten and Head Start programs and the ECERS-R for preschool classrooms in all other program types, although both measures were used primarily for self-assessment. Massachusetts used both the ECERS-R and the CLASS but the CLASS was only used as a self-assessment tool. One state only used the CLASS, and two states only used the ECERS-R. One state (Ohio) used a state-developed observation measure.

Figure III.4. Type of observation measure used in TQRIS to assess preschool classrooms in Round 1 RTT-ELC states

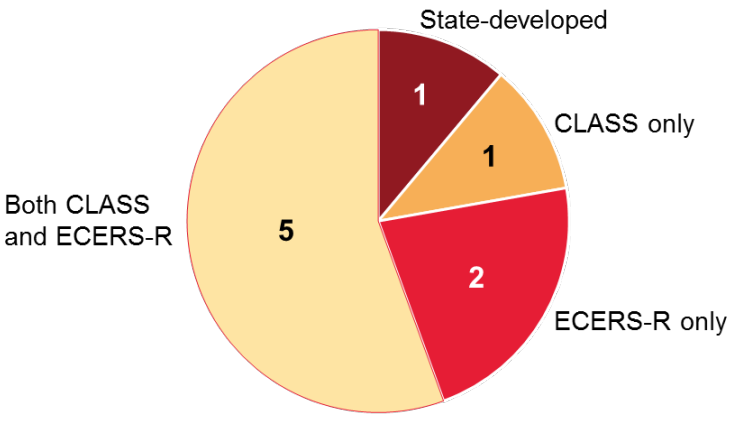


Exhibit reads: Five states used both the CLASS and the ECERS-R as classroom observation measures.
Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.
CLASS = Classroom Assessment Scoring System; ECERS-R = Early Childhood Environment Rating Scale, Revised

The observational measures were used in calculating the rating for some level(s) in eight states, while in one state the observational measures were used only to guide program improvement. In six of the eight states that used an observational measure, states averaged the scores across each observed classroom to calculate the overall score for the program and required programs to meet a minimum score to achieve specific rating levels. In addition to meeting a minimum program average, one of the six states also required a minimum age group average and three required a minimum score for each individual classroom. In a seventh state, classroom observation scores were not averaged; instead, a program's overall observation score was determined in full by its lowest-rated classroom. In an eighth state, an overall average was not calculated but each classroom had to meet the minimum score. In the one state that did not use observation scores in the rating, the state required any program in which one or more classrooms failed to meet a minimum set score to develop an improvement plan.

This page has been left blank for double-sided copying.

IV. GENERATING TQRIS RATINGS

For TQRIS to provide meaningful information on program quality, ratings must be based on accurate information and calculated consistently across programs. Previous chapters described variations in whether programs had to go through the full TQRIS data collection and verification process; for example, alternative pathways allowed certain types of programs to meet components or achieve ratings without TQRIS data collection. In this chapter, we discuss the processes followed by the Round 1 RTT-ELC states to generate TQRIS ratings, including methods for data collection and verification of components when they are assessed, the qualifications and preparation of staff involved in the rating process, and procedures for calculating the overall rating level.

A. Data collection and verification of standards

The nine Round 1 RTT-ELC states used information from various sources to verify that programs met TQRIS standards. These sources included document reviews, state databases, and interviews or observations conducted at program sites. States most often collected information from documents submitted by the programs as a starting point. For each of the 13 TQRIS components we examined, between 2 to 8 states used document reviews to gather information about a program's ability to meet standards (data not shown). Documents could include lesson plans to demonstrate curriculum requirements, assessment tools or examples of individual child reports to demonstrate child assessment, or training and credential certificates to demonstrate workforce qualifications and credentials. Programs submitted documents to the TQRIS or raters reviewed documents on-site. Document reviews were often verified or supplemented through other data collection methods such as observations or interviews, but they remained the only source of information in at least a few states for 11 of the 13 components (Figure IV.1).

For two TQRIS components—child care licensing, and workforce qualifications and credentials—states could rely on a database (outside the TQRIS system) to verify that standards were met. For example, in four states, the child care licensing database either provided a direct feed of information into the TQRIS database or could be accessed by raters to verify a program's licensing status. In Minnesota, TQRIS staff relied on a professional development registry to gather and verify information about workforce qualifications and credentials. Six additional states used professional development registries to supplement information gained from document reviews. The sophistication of professional development registries and the quality of information contained within them may vary by state, so although it is likely that using information from these systems increased the reliability of the associated TQRIS components, that is not necessarily the case for every state. States also used databases to collect information on accreditation status (in Delaware), child assessment (in Massachusetts), and provisions for children with special needs (in Massachusetts and Minnesota). One state (Massachusetts) used a database along with observation to supplement information gained from document reviews about cultural and linguistic diversity.

Figure IV.1. Sources of information for TQRIS components in Round 1 RTT-ELC states

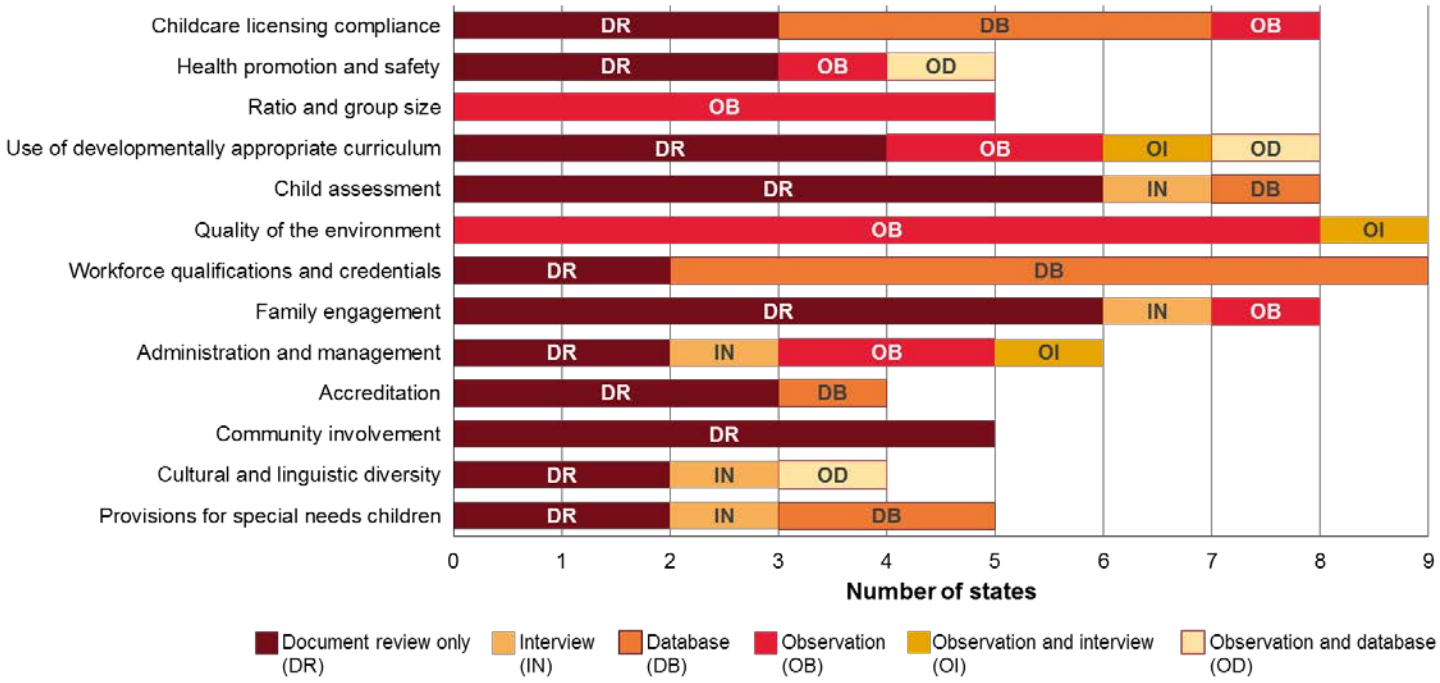


Exhibit reads: To collect information on the child care licensing compliance component of the TQRIS, four states used databases, three states conducted document reviews, and one state conducted observations. For the health promotion and safety component, five states conducted document reviews and among these five states, one also conducted observations and another supplemented document reviews with both observation and information from a database.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

[a] Information on child care licensing verification by local consortia in California was not available.

All nine states used on-site observations to collect and verify information on the learning environment. One state conducted interviews in addition to the observations. Observation measures used in classrooms serving children ages 3 to 5 included the CLASS and ECERS-R (as discussed in the prior chapter).[8] North Carolina was the only state that used on-site observations to collect information about licensing compliance. States also relied on on-site observations to supplement information collected from other sources about other components. For example, states used observations (alone or in combination with data from an existing database or interviews) to verify that programs met health and safety standards (in two states), ratio and group sizes (in five states), curriculum requirements (in four states), and family engagement standards (in one state). Three states—Delaware, Massachusetts, and North Carolina—used observations (alone or in combination with on-site interviews with program directors) to gather information about administration and management standards. Just four of the eight states that

[8] The ECERS-R is one of the Environment Rating Scales (ERS) family of measures. The appropriate ERS measure is selected based on the age of children served in the classroom observed. ERS measures other than the ECERS-R that states used to determine TQRIS ratings in center-based settings include the Infant/Toddler Environment Rating Scale, Revised Edition (ITERS-R) and the School-Age Care Environment Rating Scale (SACERS).

assessed the curriculum component used on-site observations to verify documentation submitted by programs.

In all nine states, state-certified assessors conducted the classroom observations (the CLASS and ERS) that were used to determine ratings. Three of the nine states—California, Maryland, and Rhode Island—accepted classroom observation scores that were collected for other programs in place of TQRIS-directed observations under certain circumstances. The observations had to have been completed by an external assessor, typically within the prior 12 or 13 months. All three states accepted CLASS scores generated by Head Start programs. Rhode Island accepted CLASS observations from public prekindergarten, NAEYC accredited, or Head Start programs only to meet standards for the lower rating levels; only observations conducted through the TQRIS could be used for a rating of level 3 or higher. Consortia that administered TQRIS in California could accept either the CLASS or the ERS from a range of programs but the observations from either measure had to apply to the lead teachers, child age groups, and physical location that were relevant to the child care center's participation in the TQRIS.

No state conducted observations of every classroom within a program (except in Massachusetts, for very small programs). All nine states required observations for multiple classrooms within each program and randomly selected classrooms for observation. All nine states required observations of a specified proportion of classrooms, typically one-third of the classrooms within each age group. In five states, the observations were scheduled in advance (Table IV.1). In three other states, assessors gave programs a window of time to expect them, while in Ohio the observations were unannounced.

Table IV.1. TQRIS classroom observation practices in Round 1 RTT-ELC states

State	Scheduled in advance	Minimum number of classrooms observed
California	Yes	One-third of the classrooms in each age group, with at least one in each age group
Delaware	Yes	One-third of the classrooms in each age group, with at least one in each age group
Maryland	Yes	One-half of all classrooms, with at least one in each age group
Massachusetts	Yes	All classrooms if fewer than four total; or all infant/toddler classrooms and 60 percent of all other classrooms if four or more total
Minnesota	Yes	One-third of all classrooms that had at least half the children age 33 months or older
North Carolina	No, given four-week range	One-third of the classrooms in each age group, with at least one in each age group
Ohio	No, unannounced visit	One-half of the classrooms in each age group, with at least one in each age group
Rhode Island	No, given two-week range	One-third of the classrooms in each age group, with at least one in each age group
Washington	No, given two-month range	One-third of the classrooms in each age group for the ERS and two-thirds for the CLASS, with at least one in each age group

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

B. Roles and qualifications of staff involved in the rating process

1. Staff roles

Two types of staff roles were key in the rating process: raters and assessors. Raters gathered and reviewed evidence from programs for individual TQRIS components other than learning environment; assessors were specifically focused on conducting standardized classroom observations. Raters worked within the TQRIS office and on-site at programs to conduct document reviews, interviews, and observations. Raters generally held the primary responsibility for determining a program's rating level. In five states, the assessor role was distinct from the rater role, and raters did not conduct classroom observations. In Ohio and Washington, state employees usually fulfilled the combined role of rater and assessor. In two additional states (California and Rhode Island), raters could also conduct classroom observations, but they did not typically perform both roles for the same program.

2. Staff qualifications

States varied in the qualifications required for TQRIS raters and assessors, but seven states required at least a bachelor's degree for raters and six states required at least a bachelor's degree for assessors. Four states required a bachelor's degree with a concentration in early childhood education or related field preferred for both raters and assessors (Figure IV.2). Two states required a bachelor's degree in early childhood or a related field for raters, and one of these states (Maryland), also required this degree for assessors. Ohio required raters and assessors to have a bachelor's degree or a master's degree in early childhood education or a related field, depending on the agency hiring the staff member. States made exceptions to the education requirements for raters based on years of experience (Massachusetts, Maryland and Ohio) or for individuals with desired language skills, such as fluency in Spanish (Washington). One state (Minnesota) had no education requirements for either raters or assessors. In two other states, there were no education requirements for assessors (California) or they were not set by the state (Massachusetts). In California, local county consortia set the requirements only for raters.

Figure IV.2. Minimum education requirements for TQRIS raters and assessors in Round 1 RTT-ELC states

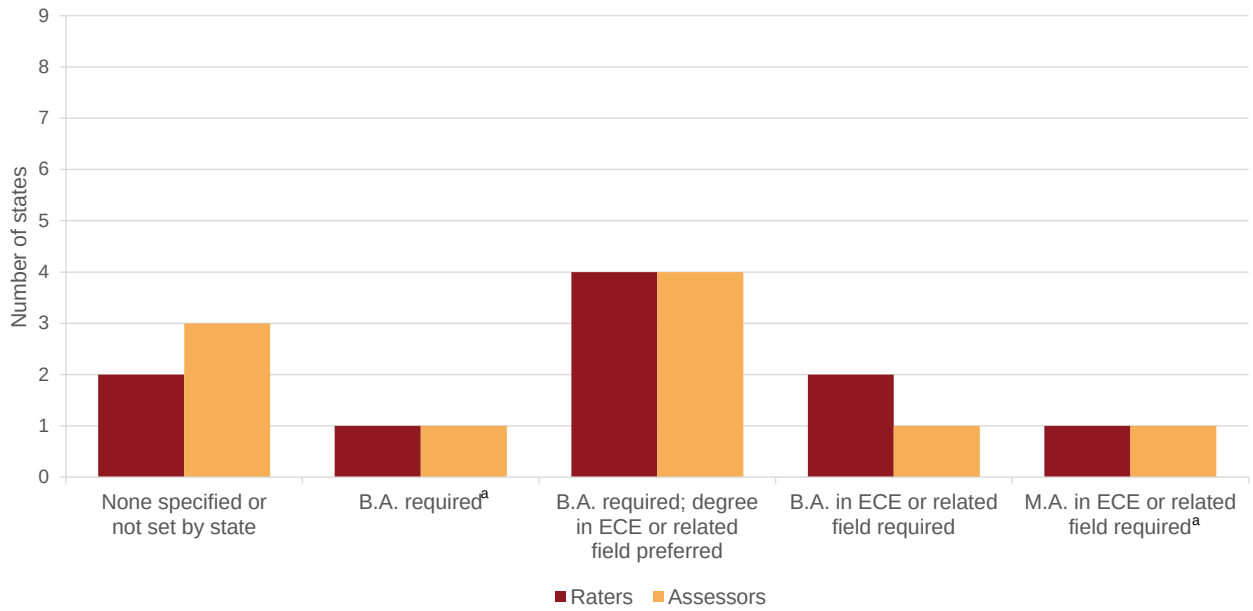


Exhibit reads: Two states did not specify education requirements for raters. Three states did not specify education requirements for assessors.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

B.A. = Bachelor of Arts degree; ECE = early childhood education; M.A. = Master of Arts degree.

[a] Ohio is counted twice because two different agencies hired state employees that fulfilled the combined role of rater and assessor, and each agency had different education requirements.

Seven states required both raters and assessors to have professional experience in education (Figure IV.3), typically two or three years in an equivalent position. Five states required raters and assessors to have experience in early childhood education. Delaware had the most demanding work experience standard of 3 to 5 years of professional experience in early childhood education for both raters and assessors. Six states additionally required assessors to have experience using classroom observation measures, but the specificity of that requirement varied. Three states required general experience or working knowledge of the measure, while the three others required that assessors meet specific training and reliability requirements to be hired. For example, in Massachusetts, ERS assessors had to be ERS certified, trained within the past five years, and have experience administering all ERS measures. Two states also required both raters and assessors to have research experience as part of a large-scale research project (Washington) or from data collection and program assessment or evaluation (Rhode Island).

Figure IV.3. Experience requirements for TQRIS raters and assessors in Round 1 RTT-ELC states

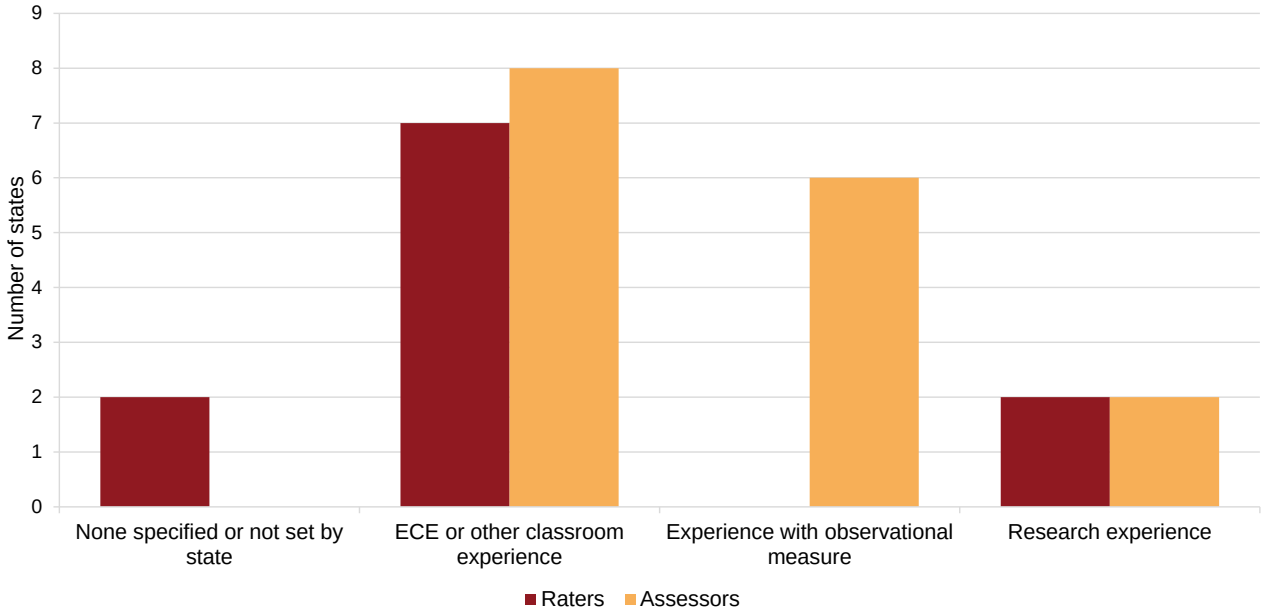


Exhibit reads: Seven states required early childhood education or other classroom experience for raters. Eight states required early childhood education or other classroom experience for assessors.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

ECE = early childhood education.

C. Training and reliability of staff involved in the rating process

Overall, the Round 1 states understood the importance of training and ensuring the reliability of staff who collected information for the TQRIS ratings. To prepare assessors, states used the training and reliability standards of the observation measure developers. States had to develop their own processes and standards to train and ensure reliability among raters, possibly explaining some of the variation in training and reliability standards for raters across states. All but two states had more formal training processes and more specific reliability thresholds for assessors than for raters.

1. Establishing reliability of assessors

Developers of classroom observation measures set training and reliability standards for their use. All nine Round 1 RTT-ELC states followed these standards and required formal training on classroom observation measures for assessors and established reliability thresholds that assessors must meet (Table IV.2). Assessors had to attend an initial in-person training conducted by the measure developer or a master trainer who had shown consistent reliability on the observation measure (referred to as an "anchor").

Table IV.2. Training and reliability requirements for classroom observation assessors, by measure used

	Environment Rating Scales (ERS)	Classroom Observation Assessment System (CLASS)	State-developed
States that used measure	California, Delaware, Massachusetts, Maryland, North Carolina, Rhode Island, Washington (7 states)	California, Maryland, Minnesota, Rhode Island, Washington (5 states)	Ohio (1 state)
Initial training requirements	Attend training by master trainer[a]	Attend training by master trainer[a]	Attend training by master trainer[a]
Initial reliability requirements	6 states: 85 percent adjacent agreement between rater and anchor[b] 1 state: 90 percent adjacent agreement between rater and anchor (Washington)	80 percent adjacent agreement between rater and anchor[b] across all scores and on 2 of 5 scores within each quality element	Exact agreement between rater and anchor[b] on 85 percent of all elements and 100 percent agreement on overall score across 3 consecutive observations
Ongoing training or reliability requirements	3 states: Reliability checks based on number of visits conducted (after every 6-10 visits in North Carolina; after every 11 visits in Washington; after every 10-20 visits in Rhode Island) 3 states: Reliability checks based on elapsed time (after every 5 weeks to 6 months in Delaware; quarterly in Massachusetts; yearly in California) 1 state: No established requirement (Maryland)	3 states: Yearly renewal trainings (California, Minnesota, Washington) 2 states: No established requirements (Maryland, Rhode Island)	Random reliability checks twice a year

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

Note: Requirements apply to all states that used the measure when the number of states is not specifically noted.

[a] A master trainer may be the developer or an individual trained by the developer who has shown consistent reliability on the observation measure.

[b] Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors. Adjacent agreement means that an assessor's score is within one point of an anchor's score and exact agreement means that an assessor's score is identical to the anchor's score.

The seven states that used the ERS set reliability standards for assessors that met or exceeded those of the measure developer. Six of the seven states required each assessor to demonstrate adjacent agreement (within one point on a 7-point scale) in scoring items the same as an anchor on 85 percent of items (typically tested over at least three visits).[9] In one other state (Massachusetts), ERS assessors had to meet a higher standard of 90 percent agreement over three visits.

The five states that had assessors administering the CLASS (one state only used it as a self-assessment) required assessors to receive training either directly from the measure developer or from state trainers. To be certified on the CLASS, assessors had to demonstrate adjacent

[9] Rhode Island did not use anchors to establish assessor reliability, instead they used the measure developer. In Washington, assessors must meet an 80 percent initial reliability on the CLASS based on one online training.

agreement (within one point on a 7-point scale) on 80 percent of assigned scores and on 2 of 5 scores within each of 10 quality dimensions.

Ohio did not use either the ERS or the CLASS but required a reliability level of 85 percent over at least three observations for assessors using their state-developed measure. That is, an assessor had to demonstrate exact agreement with the anchor on the scores for each of the elements 85 percent of the time for three consecutive observations. They also had to agree on the overall score 100 percent of the time across the three observations.

2. Maintaining reliability of assessors

Once initial reliability is established, assessors are required to complete ongoing reliability checks. The frequency of ongoing checks varied among the seven states using the ERS. Three states conducted ongoing reliability checks based on the number of visits, ranging from every 6 visits to every 20 visits, depending on the assessor's experience. Three other states conducted ongoing reliability checks based on elapsed time, ranging from every five to seven weeks up to once a year. The seventh state (Maryland) did not yet have an established procedure for maintaining ERS reliability. Three of the five states that used the CLASS required assessors to attend yearly "renewal" trainings through the measure developer. The other two states did not yet have established procedures for maintaining assessors' reliability on the CLASS. In Ohio, assessors underwent random reliability checks twice a year.

3. Establishing reliability of raters

States determined their own processes for training and ensuring reliability among raters (staff who gathered and reviewed evidence from programs for individual TQRIS components other than learning environment). California does not have a statewide protocol for raters because training protocols are established by local consortia. Eight states provided information on the process used to train raters and seven had processes in place, though mostly informal, to establish initial reliability in assigning ratings (Table IV.3). Training included individual or group sessions covering topics such as the goals and purposes of the TQRIS, TQRIS standards and licensing requirements, and procedures for reviewing evidence. Five states required new raters to shadow experienced raters before they rated programs on their own. Five states provided raters with protocols or other procedural documents. Three states provided new raters with opportunities to practice the rating process through sample documents or exercises that simulated different rating decisions. One state (Maryland) connected all new raters with a mentor-rater to serve as a resource.

Two states specified an initial reliability threshold that raters had to meet before conducting ratings independently. In Minnesota, raters had to reach 85 percent agreement in scoring with an anchor across five programs before they could begin rating independently. In Ohio, raters had to demonstrate 85 percent agreement with an anchor on rating elements across three consecutively rated programs and achieve 100 percent agreement on the overall met/not met score for a program.

Table IV.3. Training and reliability requirements and procedures for raters in Round 1 RTT-ELC states

	Total number of states with requirements or procedures	Requirements and procedures	Number of states
Initial training and reliability	8	Individual or group training sessions	8
		Shadow current raters	5
		Procedural documents	5
		Practice rating process	3
		Assignment to a mentor	1
		Establish reliability with anchor	2
Ongoing reliability	8	Group meetings, informal communication	3
		Double coding by two raters	2
		Ongoing reliability checks at specific frequency	2
		Periodic audits	1

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

4. Maintaining reliability of raters

Processes for ensuring ongoing reliability of raters were less well-established compared with procedures for assessors, but eight states had some procedures in place. There were no statewide protocols for ensuring reliability in California; instead, training protocols were determined by local consortia. Three states relied on informal processes to maintain consistency in scoring across raters, such as regular team meetings and informal communication. Two states ensured inter-rater reliability by requiring two raters to review the same records and work to reach consensus (referred to as double-coding). Two states conducted ongoing reliability checks between raters and an anchor at a specific frequency. In Ohio, reliability checks occurred twice a year and in Minnesota reliability was checked after every 10th rating. In Maryland, supervisors reviewed ratings for randomly selected cases, and the state required a minimum of 80 percent agreement in the supervisor's and supervisee's ratings. Raters that did not meet this requirement entered into a probationary period with more frequent review and a 90 percent agreement requirement. Raters who were unable to meet requirements during the probationary period were terminated.

D. Calculating ratings

The many steps in data collection, verification, and calculation to determine final TQRIS ratings can make the ratings vulnerable to errors, and automating these steps can increase the reliability of the process. All states, with the exception of California and Rhode Island, electronically stored the data used to generate TQRIS ratings in an online database at the state level (Table IV.4).[10] Raters or data entry staff entered the information in states where it was not collected electronically. Five states automated the process of calculating component scores or

[10] TQRIS data were maintained by local consortia in California.

tracking whether TQRIS components were met, depending on the rating structure. These same five states also automated the determination of the final rating level.

Table IV.4. Automated TQRIS data collection and calculations in Round 1 RTT-ELC states

	States
Electronic data on assessed components or indicators stored in online database at state level	All states except California and Rhode Island[a]
Automated system to calculate points or track that components are met	California, Maryland, Massachusetts, Minnesota, North Carolina
Automated calculation or assignment of final rating	California, Maryland, Massachusetts, Minnesota, North Carolina

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

[a] A TQRIS data system was under development in Rhode Island. All states used indicator-level data in calculating ratings but not all states calculated component-level scores.

V. TQRIS REFINEMENTS SINCE RECEIPT OF RTT-ELC GRANT

RTT-ELC funds were intended, in part, to support the development and implementation of TQRIS. Given that all but two of the nine Round 1 RTT-ELC states were already implementing a TQRIS to some degree prior to the grant award, much of what occurred over the grant period was continued growth and additional refinements to the existing systems. Prior chapters of this report described the status of TQRIS structures and policies at the end of 2014 and early 2015. The report documents the variation that existed across the many dimensions of TQRIS, suggesting that TQRIS are still in an exploratory and dynamic stage of development.

Since receiving RTT-ELC grants in 2011, the nine Round 1 RTT-ELC states all modified their TQRIS in some way. Most commonly, the states refined TQRIS standards and rating processes. Fewer revised the eligibility criteria for TQRIS participation, data collection and verification processes, the use of alternative pathways, and the provision of supports such as technical assistance and participation incentives (Figure V.1).

To illustrate which areas were in flux during the RTT-ELC grant period, this chapter summarizes the refinements that states made to their TQRIS since receiving the grants and why. The study team collected information about refinements made in all nine states; however, it was beyond the scope of the study to collect details about the processes that states used to make these refinements, including who made the decisions and how. (The information presented in earlier chapters reflects the status of each TQRIS after refinements were made, except where ongoing or future refinements are described.)

Figure V.1. Refinements to TQRIS made by Round 1 RTT-ELC states since receipt of RTT-ELC grant

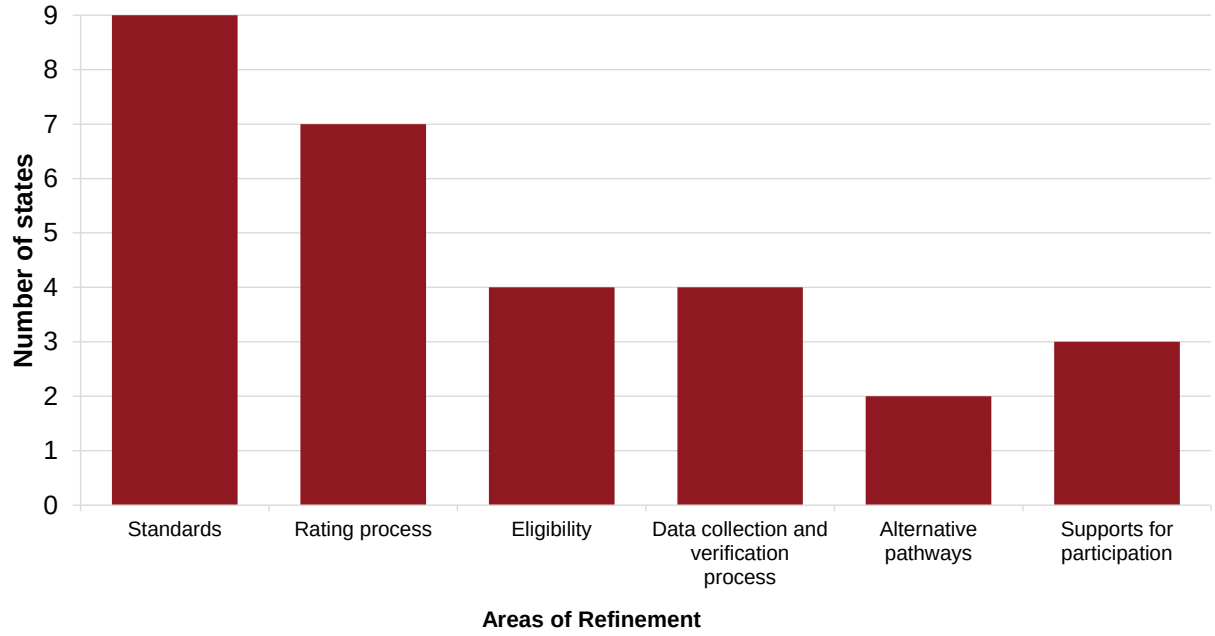


Exhibit reads: Nine states made refinements to TQRIS standards after receiving their RTT-ELC grant.
Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.

A. Refinements to TQRIS standards

All nine states refined TQRIS standards—whether major changes (such as adding or eliminating a standard) or minor clarifications (such as shifting the focus of a standard from the program to the classroom level) since their receipt of the RTT-ELC grants. Four states also adjusted the rating level at which certain standards applied or the number of points assigned for meeting a standard. States revised standards according to stakeholder feedback and evaluation findings. For example, in Maryland, the language used in standards was refined based on focus group feedback and online survey responses from participating programs. In Minnesota, one of the TQRIS rating components was dropped after findings from a pilot study revealed limited variation. Rhode Island made changes that aligned standards across child care licensing, public prekindergarten, and TQRIS.

TQRIS continues to evolve, and three states are currently making additional revisions to standards. For example, one state (Delaware) with a point-based structure is phasing in six essential standards that programs will be required to meet at the two highest rating levels because of stakeholder interest in making certain standards non-negotiable. Programs will still be able to earn points for other standards they pursue, but the six essential standards will no longer be optional.

B. Refinements to the rating process

Seven states modified the process for determining TQRIS rating levels, such as the rating structure, or the use of observational measures and how they are scored.

1. Rating structure

Four states overhauled the rating structure, primarily to offer programs increased flexibility in how they demonstrated quality. Three states gave up the building-block approach, in which all standards at a given level had to be met to achieve a particular rating. They adopted either a points or hybrid system in which programs chose, at least in part, which standards to pursue as they worked toward a particular rating. Conversely, one state (Minnesota) eliminated some degree of choice in the standards that programs could pursue by moving from a points system to a hybrid model after the TQRIS pilot determined that ratings were highly concentrated at the high end of the scale.

2. Observational measures

Six states revised their use or scoring of observational measures within TQRIS. Three of the states changed the way they used classroom observation measures by introducing their own, newly created measure, switching from the ERS to the CLASS, or eliminating a subscale from an existing measure. The other three states decreased the role of observational measure scores in determining a rating level. One state fully eliminated an observational measure from the ratings determination and used the measure solely to guide program improvement. Another state moved to averaging classroom observational scores to determine a program's rating level and no longer required each classroom to meet a minimum score. The third state began using the ERS developer's alternative method for scoring in order to reduce the impact of a few indicators on subscale scores.

C. Additional refinements

Four or fewer states made refinements to expand eligibility, revise data collection and verification processes, alter the use of alternative pathways to achieving higher rating levels or increase the supports provided to programs.

1. Eligibility criteria

To increase the participation of programs serving low income children, four states expanded TQRIS eligibility to school-sponsored prekindergarten programs or made enrollment in TQRIS a requirement for receiving child care subsidy payments. One of the four states (Washington) also extended TQRIS eligibility to Head Start programs.

2. Data collection and verification

Four states changed their data collection and verification processes by clarifying documentation requirements or revising the timing of classroom observations. Two states took opposite approaches to the timing of observational measures. One changed to first requiring classroom observations before verifying other TQRIS components because a minimum score was required for each rating level. The other changed to first requiring a review of submitted documentation to determine if a program would meet the minimum requirements for the rating level at which an observation was required.

3. Accelerated pathways

To promote participation in TQRIS, Washington added an accelerated pathway to upper-level ratings for public prekindergarten and Head Start programs. However, Minnesota, which had automatically granted public prekindergarten and Head Start programs the top rating, decided to make a shift to an alternative pathway that waived many TQRIS standards but still required these programs to meet standards related to curriculum and the use of child assessments.

4. Technical assistance and financial supports for TQRIS participation

Three states increased the level of technical assistance provided to programs by extending coaching support to programs at the lower rating levels, implementing monitoring visits to identify areas in need of program improvement, and offering additional supports to help programs meet forthcoming standards. One of these three states (Delaware) also increased the child care subsidy rates the state paid to programs that achieved the two highest rating levels.

This page has been left blank for double-sided copying.

VI. DISCUSSION

This report focused on the progress made and challenges faced by the nine Round 1 RTT-ELC states in achieving the first three objectives for TQRIS under the RTT-ELC grants: (1) developing and adopting a common, statewide TQRIS, (2) promoting participation in the TQRIS, and (3) rating and monitoring early learning and development programs. Previous chapters presented key findings related to these objectives.

In this chapter, we first present a key finding that cuts across all three of these state objectives. We then turn to the fourth and fifth objectives—(4) promoting access to high quality programs for children with high needs, and (5) validating the effectiveness of the TQRIS. Although these objectives were not the focus of this report, we mention the other studies that are being conducted to shed light on whether states achieved these objectives, and we present our findings on key features of states' TQRIS that may affect their ability to achieve these objectives. We conclude with a summary of the progress and challenges in developing TQRIS in the nine Round 1 RTT-ELC states.

There is tremendous variation across states in TQRIS structure, features, and processes; each state's TQRIS is uniquely designed and implemented. This key finding cuts across all three state objectives that were the focus of this report. States differ in terms of the timing of implementation, the policies used to promote participation, the amount and type of data available about programs and the children enrolled in them, the methods used to classify programs by type, the rating structure, the number of TQRIS components that contribute to the final rating, the way components are measured and defined, and how components are combined to arrive at the final rating.

A. Promoting access to high quality programs for children with high needs

Additional data are needed to fully understand the distribution of programs across the rating levels, increases in the number of programs in the top rating levels of TQRIS, and the number of children with high needs enrolled in top-rated programs. States were required to report their progress on the five RTT-ELC objectives to the U.S. Department of Education each year. Each state report included data on the distribution of TQRIS participating programs across the rating levels, and on two performance measures specific to this RTT-ELC objective: (1) increasing the number of early learning and development programs in the top tiers of the TQRIS, and (2) increasing the number and percentage of children with high needs who are enrolled in early learning and development programs that are in the top tiers of the TQRIS. These data are difficult to assess across the states given the differences in data coverage of children and programs, the methods of classifying programs by type, and the methods of counting programs in rating levels. In addition, because these data represent a single point in time (that is, they are not longitudinal data), it is challenging to discern patterns of TQRIS entry and movement up the rating levels by specific program types over time.

As a follow-up to this descriptive study, the study team is working with administrators in the nine states to collect data on TQRIS participation and ratings from 2012 through 2016 to examine the extent to which participating programs moved from lower to higher rating levels. In

addition, we are collecting information about all programs in the state—both those that participated in the TQRIS and those that did not—to provide a more complete picture of how access to high-quality programs may have changed under RTT-ELC. For example, improvements in the rating levels of programs that serve children with high needs may be a better indicator of a state's ability to improve access to high quality programs than increased TQRIS participation of programs that may have already been of high quality before participating in the TQRIS.

B. Validating the effectiveness of the TQRIS

The nine Round 1 RTT-ELC states have all commissioned independent studies to validate the effectiveness of their TQRIS; these studies are expected to be completed in 2017. However, through this descriptive study of the nine states, we identified some key features of states' TQRIS that may affect the reliability and validity of their ratings.

States faced trade-offs between promoting participation and quality improvement and producing valid, reliable ratings. These trade-offs are reflected in the structure of TQRIS in ways that can make it difficult to assess how accurately the ratings reflect program quality. For example, alternative pathways are intended to ease access and reduce the burden of participation among programs that have already met similar standards to those of the TQRIS. However, the underlying assumption that alternative pathway programs are comparable in quality to TQRIS rated programs at the same level has not been empirically confirmed.

TQRIS ratings are structured as a lattice to bring programs into the system (typically by meeting licensing standards) and to promote quality improvement by giving programs some flexibility in how they demonstrate quality (for example, by allowing programs to apply for a certain rating level or to select specific TQRIS components for which to receive points). TQRIS ratings are not structured as standardized scales on which all programs are assessed for each component in the same way.

In addition, the reliability standards to which assessors who conducted observations were held were not formulated by measure developers with TQRIS in mind. For example, in many cases these standards required assessors to achieve only adjacent rather than exact agreement with the ratings of an anchor. These standards may allow for more inconsistency in measurement than is desirable in assessments that are used for high-stakes purposes, such as when a single point difference on a seven point scale can result in a higher quality rating.

The complex composition of TQRIS ratings may weaken the relationship between ratings and child outcomes. TQRIS ratings in the nine states are based on many components. Five of the nine states included 10 or more of the 13 components examined in this report. Only 4 of the 13 components (child-to-staff ratios and group size, curriculum, staff qualifications, and environment) have been associated with child outcomes in the research literature. Although the evidence suggests that associations are small, these components are included in most of the nine TQRIS: all nine systems included the latter two components, seven included curriculum, and five included ratios and group size (presumably because ratios and group size are also set in licensing standards). However, the nine states also included other components, for example, to emphasize administration and management, family engagement, and community involvement. These

components are less directly related to children's classroom experiences, which may further weaken the relationship between TQRIS ratings and child outcomes.

In addition, the extent to which TQRIS ratings reflect true underlying differences in program quality depends on the richness of the measures used to assign them. Some standards are challenging to operationalize, much less measure. For example, to earn certain ratings, eight states required programs to use developmentally appropriate curricula. The states typically checked this standard based on a program self-report. However, there was likely substantial variation in the use of curricula between programs that a simple yes/no indicator cannot capture, and states did not have the resources to extensively assess each center's curriculum. Also, the inclusion of measures that add complexity but do not capture much meaningful variation in quality may contribute to weak relationships between ratings and child outcomes. State validation studies could help address this by identifying the TQRIS rating components associated with child outcomes (and at what thresholds) and ways to simplify the construction of ratings.

C. Progress and challenges in developing TQRIS

The nine Round 1 RTT-ELC states worked to refine their systems to promote TQRIS participation among programs (particularly those that serve low-income children), define standards, and improve data collection and verification—with the goal of building systems that signal and promote quality. The nine states made progress in structuring policies to try to encourage a range of early learning and development programs to participate and improve their quality rating. The states have also put policies and practices in place to collect information, hire and train rating staff, and calculate ratings, which may ultimately help to reliably assess program quality. The states' validation studies, expected to be completed in 2017, may shed further light on the ability of TQRIS ratings to convey meaningful differences in program quality.

The study team is working with the nine states in 2016 and 2017 to gather administrative data that we will use, together with the descriptive information included in this report, to examine patterns in TQRIS participation and ratings during the RTT-ELC grant period. In particular, this follow-up analysis will address two questions: (1) what are the patterns in TQRIS participation, the distribution of programs and children across the rating levels, and the movement of programs and children across the rating levels since 2011?, and (2) how do TQRIS characteristics (including structure, policies, and practices) relate to program movement up the rating levels and achievement of the highest level? A systematic and consistent analysis that captures entry and movement of programs over time across the states will explore whether changes in the number of programs at the highest rating levels are driven by movement of programs from low to high-quality ratings or increased TQRIS entry of high-quality programs. In addition, the analysis will shed light on program responses to TQRIS policies by examining how participation and ratings relate to different policies across states.

This page has been left blank for double-sided copying.

REFERENCES

Applications for New Awards; Race to the Top—Early Learning Challenge, 2011. Available at http://www.gpo.gov/fdsys/pkg/FR-2011-08-26/pdf/2011-21756.pdf. Accessed April 9, 2015.

Burchinal, M., N. Vandergrift, R. Pianta, and A. Mashburn. ."Threshold Analysis of Association between Child Care Quality and Child Outcomes for Low-Income Children in Prekindergarten Programs." *Early Childhood Research Quarterly*, vol. 25, 2010, pp. 166-176.

Caronongan, P., G. Kirby, K. Boller, E. Modlin, J. Lyskawa. "Assessing the Implementation and Cost of High Quality Early Care and Education: A Review of Literature." OPRE Report 2016-31.Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, 2016.

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

Kirby, Gretchen, Pia Caronongan, Lizabeth Malone, and Kimberly Boller. "What Do Quality Levels Mean? Examining the Implementation of QRIS Ratings to Inform Validation." *Early Childhood Research Quarterly*, vol. 30, 2015, pp. 291-305.

Mathematica Policy Research. "Race to the Top-Early Learning Challenge (RTT-ELC): Descriptive Study of Tiered Quality Rating and Improvement Systems (TQRIS) Master Data Collection Protocol, January 2015." Princeton, NJ: Mathematica Policy Research, 2017. Available at https://www.mathematica-mpr.com/-/media/publications/pdfs/RTT ELC Master Data Collection Protocol.pdf. Accessed June 27, 2017.

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.

Pianta, Robert C., Karen M. La Paro, and Bridget K. Hamre. "CLASS Classroom Assessment Scoring System Manual." Baltimore, MD: Brookes Publishing, 2008.

QRIS National Learning Network. "State Resources and Profiles." Available at http://qrisnetwork.org/state-qris-resources. Accessed January 15, 2014.

The BUILD Initiative and Child Trends. "A Catalog and Comparison of Quality Rating and Improvement Systems (QRIS)", 2016. Available at http://qriscompendium.org/. Accessed April 12, 2017.

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, 2010.

U.S. Department of Education, Race to the Top – Early Learning Challenge program, Awards. Available at https://www2.ed.gov/programs/racetothetop-earlylearningchallenge/awards.html. Accessed November 15, 2014.

Yoshikawa, H., C. Weiland, J. Brooks-Gunn, M. Burchinal, L. Espinosa, W. Gormley, W. J. Ludwig, K. Magnuson, D. Phillips, and M. Zaslow. "Investing in Our Future: The Evidence Base on Preschool Education." New York: Foundation for Child Development and Ann Arbor, MI: Society for Research in Child Development, 2013.

Zaslow, M., R. Anderson, Z. Redd, J. Wessel, L. Tarullo, and M. Burchinal. "Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature." OPRE 2011-5. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2010.

Zellman, G.L., R.N. Brandon, K. Boller, and J.L. Kreader. "Effective Evaluation of Quality Rating and Improvement 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, 2011.

Zellman, G.L., and R. Fiene. "Validation of Quality Rating and Improvement Systems for Early Care and Education and School-Age Care." Research-to-Policy, Research-to-Practice Brief OPRE 2012-29. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2012.

Zellman, G.L., M. Perlman, V. Le, and C. Setodji. "Assessing the Validity of the Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child-Care Quality." Pittsburgh, PA: RAND Corporation, 2008.

APPENDIX A

PROFILES OF TIERED QUALITY RATING AND IMPROVEMENT SYSTEMS (TQRIS) IN THE NINE ROUND 1 RACE TO THE TOP-EARLY LEARNING CHALLENGE (RTT-ELC) STATES

This page has been left blank for double-sided copying.

California Race to the Top – Early Learning Challenge Grant Tiered Quality Rating and Improvement System

Grant details

Grantee: California Department of Education

TQRIS lead agency: California Department of Education; California's TQRIS comprises 17 lead consortia agencies in 16 counties

RTT-ELC grant amount: $75,000,000

TQRIS implementation status

Pilot: No formal pilot

Implementation date: 2012 (16 counties); 2013 (14 additional mentee counties)

Date of current TQRIS standards: September 2012

TQRIS structure[a]

Defining rating levels: Hybrid system with five levels. Level 1 uses a block structure. Levels 3 and 4 use a point structure; the structure of levels 2 and 5 is locally determined. Levels 1, 3, and 4 use the same requirements across all consortia. The minimum requirements for levels 2 and 5 are the same for all consortia, but local consortia can add requirements.

Entry into and progression through TQRIS:

- A license (or licensing compliance) in good standing is the only requirement at level 1.
- Based on points earned, programs may be awarded a level 1 through 5 rating (at entry or to advance).
- Programs are not subject to a minimum or maximum time limit at any level.
- Ratings are valid for two years.

TQRIS participation

Participation requirements: Voluntary for all programs

Participation of center-based early care and education (ECE) programs (as of September 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
State-funded preschool programs	553 (18%)	27,016 (n.a.)
Head Start programs	350 (18%)	18,070 (n.a.)
Licensed, center-based programs	162 (9%)	6,125 (n.a.)
License-exempt, center-based programs	19 (20%)	195 (n.a.)

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating.

n.a. = not available

Distribution of programs and children across rating levels (as of September 2014)

	Number of center-based ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Participating but not yet rated	221	6,522
Level 1	21 (2%)	1246 (3%)
Level 2	64 (7%)	3093 (7%)
Level 3	328 (38%)	17,310 (39%)
Level 4	395 (46%)	20,224 (45%)
Level 5	55 (6%)	3,011 (7%)

Note: Program counts by level were calculated based on state-reported data on the number of programs in each level by program type. Because program type categories were determined by funding source or operating authority, programs with more than one funding stream and/or operating authority may be counted twice at a particular level.

ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- No alternative pathways

Incentives:

- The state does not provide tiered reimbursement rates.
- Local consortia use block grant funds from the state to provide quality improvement support to help state preschool providers achieve a level 4 or 5 rating, and to reward state preschool providers that have achieved a level 4 or 5 rating.
- Most consortia have developed local incentives for programs to improve their Environment Rating Scale (ERS) scores, encourage participation, and promote continuous quality improvement.

Quality components[b]

Component categories: Licensing compliance, health promotion and safety, ratio and group size, child assessment, environment, and workforce qualifications and credentials.

Observational measures: External assessors for programs in levels 3 through 5 conduct CLASS and ERS assessments.

Generating ratings

Data collection, verification, and calculation of ratings: All programs submit to local consortia a self-reported portfolio and provide common documents demonstrating level of quality. External observers visit programs to conduct observations. Consortia-certified raters calculate overall ratings. Some calculation methods vary by consortia because at levels 2 and 5 local consortia can determine the rating structure and add requirements.

Staff qualifications, training, and reliability requirements

	Program raters	Observational measure assessors
Type of staff	Consortia-certified staff	Consortia-certified staff
Required education/ qualifications	Determined by consortia	No state-specified educational requirements. Experience in early education and skills or qualifications beyond experience are required. Assessors must also be independent and reliable to each instrument.
Training/support	Determined by consortia	Assessors attend training to learn about using the scales/instrument and complete several practice observations. An author or an anchor must deliver the initial assessor training.*
Initial reliability requirement	Determined by consortia	ERS: 85 percent agreement (within one point) with the consensus scores of an anchor or level 1 observer across three consecutive ratings CLASS: Annual certification through Teachstone
Ongoing reliability requirement	Determined by consortia	ERS: Annual recalibration; assessors must have 85 percent agreement (within one point) with the consensus scores of an anchor or level 1 observer across three consecutive ratings. CLASS: Annual certification through Teachstone.

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

ERS = Environment Rating Scales; CLASS = Classroom Assessment Scoring System

Key revisions from 2011 to 2014

- Changed from a block rating structure to a point structure and ultimately implemented a hybrid structure to provide flexibility to participating programs and recognize diverse areas of quality.
- Streamlined the components on which center-based programs were evaluated, reducing from 16 to 7 components.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of center-based early care and education programs were provided by the state.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Delaware Stars for Early Success

Grant details

Grantee: Delaware Department of Education

TQRIS lead agency: Delaware Department of Education

RTT-ELC grant amount: $49,878,774

TQRIS implementation status

Pilot: 2007

Implementation date: 2008 (statewide implementation)

Date of current TQRIS standards: July 2014

TQRIS structure[a]

Defining rating levels: Point system with five levels

Entry into and progression through TQRIS:

- Licensing compliance is a prerequisite for all programs to participate with the exception of public school programs funded by Individuals with Disabilities Act (IDEA), Part B, section 619 that do not have to be licensed. Licensing compliance is the only requirement at level 1.
- Except for programs eligible for the accelerated pathways, all programs must enter at level 1. To move from level 1 to 2, programs must complete an orientation. After earning level 2, programs may apply for any of levels 3 through 5.
- Programs may spend a maximum of one year at level 1 and 2 years at level 2. If a program has not moved up from level 1 or 2 within these timeframes, they are removed from Stars and encouraged to re-apply.
- Ratings are valid for three years for levels 3 through 5.

TQRIS participation

Participation requirements: The state-funded Head Start equivalent program called the Early Childhood Assistance Program (ECAP) has made participation mandatory; it is voluntary for all other programs; license-exempt programs are not eligible. Because all Head Start programs are also ECAP programs, there is 100 percent participation among Head Start programs.

Participation of early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
Programs receiving CCDF subsidies	478 (51%)	n.a.
Licensed center-based programs not receiving CCDF subsidies	18 (4%)	n.a.
ECAP programs	11 (100%)	n.a.
Head Start and Early Head Start programs	31 (100%)	n.a.
Programs funded by IDEA, Part B, section 619	6 (38%)	n.a.
Programs funded under Title I	12 (63%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include center-based, family child care, and school-age programs.

ECAP = Early Childhood Assistance Program; IDEA = Individuals with Disabilities Act; CCDF = Child Care and Development Fund
n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	54 (11%)	n.a.
Level 2	140 (29%)	n.a.
Level 3	66 (14%)	n.a.
Level 4	150 (31%)	n.a.
Level 5	68 (14%)	n.a.

Note: Counts include center-based, family child care, and school-age programs.
ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- Head Start and ECAP programs are awarded a level 4 rating upon completion of the Head Start-specific orientation and may complete an Environment Rating Scale (ERS) assessment to move to level 5.
- Public school programs funded by Individuals with Disabilities Act (IDEA), Part B, section 619 enter at level 1 but often progress more quickly than other programs because the requirements for 619 also fulfill level 4 requirements. Programs that develop a quality improvement plan and earn the appropriate ERS cut score may earn a level 5.
- Title I programs and school-based parent-pay programs start at level 1 but automatically earn 18 and 15 points (out of a 100 possible points), respectively.
- Programs accredited by the National Association for the Education of Young Children (NAEYC) are automatically awarded a level 5 rating following completion of the NAEYC-specific Delaware Stars orientation.

Incentives:

- Tiered reimbursement for child care subsidy payments are available to providers with a rating of level 3 or higher.
- Quality improvement grants, ranging from $2,500 to $5,000, are available to participating programs to promote quality improvement. The amount varies with the size of the program.
- Programs receive merit awards, ranging from $300 to $500, for achieving a specific rating level.
- Infrastructure fund grants are available to level 2 programs and higher to support capital improvements and technology needs.
- Compensation, Retention and Education (CORE) awards are paid directly to staff in participating programs if they reach or progress past step 4 on the Delaware Early Childhood Career Lattice. T.E.A.C.H. scholarships are also available for all programs, but priority access goes to staff in programs participating in TQRIS.

Quality components[b]

Component categories: Licensing compliance, ratio and group size, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, administration and management, accreditation, community involvement, cultural and linguistic diversity, and provisions for special needs.

Observational measures: ERS assessments (parent and staff and personal care routines subscales are excluded).

Generating ratings

Data collection, verification, and calculation of ratings: Programs choosing to pursue level 3 through 5 ratings schedule an ERS observation conducted by a classroom observation assessor. A

program rater verifies that the program has met its chosen standards through document reviews, interviews, and observations. Raters assign points to each relevant indicator and then sum the points across indicators to determine a program's rating.

Staff qualifications, training, and reliability requirements

	Program raters	Classroom observation assessors
Type of staff	Contracted staff	Contracted staff
Required education/ qualifications	Bachelor of arts required, preferably in early care and education; three to five years of experience in early education.	BA required, preferably in early care and education; three to five years of experience in early education and ERS experience.
Training/support	Two raters are responsible for overseeing the rating process. A trainer serves as an anchor for the two raters. A standards verification manual guides the rating process.	Each new assessor is trained by an assessment anchor on one ERS scale at a time.*
Initial reliability requirement	n.a.	85 percent with anchor across three reliability assessments
Ongoing reliability requirement	n.a.	Assessor checks occur every five to seven weeks for a six-month period. If the checks are consistently high, assessors move to reliability checks every three months.

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

n.a. = not available; ERS = Environment Rating Scales

Key revisions from 2011 to 2014

- Switched from a building block to a point rating structure to offer programs more flexibility in how they demonstrated quality based on feedback from providers during the pilot.
- Revised specific standards and the number of points allocated to each. For example, the state added one point to the total a program could earn for the child assessment standard and specified that assessments must be completed a minimum of two times per year.
- Eliminated the use of the ERS Personal Care Routines subscale to focus on intentional teaching and school readiness rather than health and safety, which was covered by licensing.
- Revised the verification manual to clarify the rationale for each standard, describe methods for verifying that a program meets each standard, and provide explicit definitions for common terms to improve understanding of TQRIS standards and procedures.
- Switched to conducting ERS observations before raters verified other components to enhance efficiency of the rating verification process because a minimum ERS score was required for each rating level.
- Increased the Child Care and Development Fund subsidy rates paid to programs that achieved a level 4 or 5 rating.
- Began phased implementation of six essential standards that were not previously required to achieve a level 4 or 5 rating.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Maryland EXCELS

Grant details

Grantee: Maryland State Department of Education (MSDE)

TQRIS lead agency: MSDE

RTT-ELC grant amount: $49,999,143

TQRIS implementation status

Pilot: 2012–2013

Implementation date: Statewide July 2013

Date of current TQRIS standards: Child care center standards, March 2014; Public prekindergarten standards, November 2012 (currently undergoing revision)

TQRIS structure[a]

Defining rating levels: Building-block system with five levels

Entry into and progression through TQRIS:

- License or letter of compliance plus additional requirements required for level 1; license-exempt programs are eligible and do not have to become officially licensed.
- All programs may apply for a specific rating level (at entry or to progress).
- Programs may remain at each rating level for any length of time. (No minimum or maximum time requirement at any rating level.)
- Ratings are valid for one year.

TQRIS participation

Participation requirements: Participation is an eligibility requirement for public prekindergarten expansion grants, and, as of July 2015, programs will be required to enroll in Maryland EXCELS to receive Child Care and Development Fund (CCDF) subsidies; voluntary for all other programs.

Participation of center-based early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in the state)
Licensed center-based, Head Start, and license-exempt programs	1,095 (55%)	31,589 (n.a.)
School-sponsored prekindergarten	7 (1%)	n.a.
Total	1,102	n.a.

n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of center-based ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Unrated*	679	n.a.
Level 1	230 (54%)	n.a.
Level 2	70 (17%)	n.a.
Level 3	28 (7%)	n.a.
Level 4	22 (5%)	n.a.
Level 5	73 (17%)	n.a.

* In process (rating not yet determined); all school-sponsored prekindergarten programs are currently unrated because the state is in the process of finalizing standards for these programs.

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- An accelerated pathway to a level 4 or 5 rating is available to school-based public prekindergarten programs.
- Accredited programs receive credit for criteria that align with the standards of the accrediting agencies recognized by the state (the aligned criteria vary by type of accreditation). Accreditation is required to earn a level 5, but accredited programs do not automatically earn a level 5.

Incentives:

- Tiered reimbursement for child care subsidy payments is available only to providers with a rating of level 3, 4, or 5; differential payments range from 10 to 44 percent above set rates depending on the age of the child, the type of care, and the rating level.
- Financial bonuses, ranging from $50 to $4,500 (based on capacity and level), for each level achieved (from 1 to 5); bonuses to programs that maintain a level 4 or 5 may be awarded subject to funding availability.
- Bonuses to individual teachers for attainment of a Maryland Child Care Credential.
- Quality improvement grants and supports are available to support accreditation, staff professional development, and other activities related to quality improvement.

Quality components[b]

Component categories: Licensing compliance, health promotion and safety, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, administration and management, accreditation, community involvement, cultural and linguistic diversity, provisions for special needs

Observational measures: The Environmental Rating Scale (ERS) and CLASS assessments are used for program improvement purposes but are not included in the rating calculation. The development of program improvement plans is based on subscale scores. Programs must complete an ERS or CLASS self-assessment at level 3 and must have an ERS or CLASS conducted by an approved assessor at level 4. MSDE observes programs that have achieved a level 5 rating once every three years.

Generating ratings

Data collection, verification, and calculation of ratings: Programs must submit documentation to support all component requirements except for licensing. Documentation for workforce qualifications is verified through existing databases. Accredited programs are automatically credited for TQRIS criteria that align with accreditation standards. A rating is automatically calculated by the

application system and verified by the program coordinators. An MSDE quality assurance specialist reviews the rating, and then the final rating is approved by a Maryland Excels administrator.

Staff qualifications, training, and reliability requirements

	Program raters	Observational measure assessors
Type of staff	Contracted staff	Contracted staff
Required education/ qualifications	Bachelor of arts degree in early childhood education or related field required	Bachelor of arts degree in early childhood education or related field required
Required qualifications	Experience specifically in early childhood education	Knowledge of ECE programs generally and of specific measure; five years of experience in teaching/administering an ECE program
Training/support	Guidebooks and rating rubrics available; connects new raters with a mentor-rater; simulator tool to rate sample documents under development	Training by state anchor; anchor trained by tool developer*
Initial reliability	Random secondary reviews by the lead rater; must achieve 80 percent	85 percent with anchor across three visits
Ongoing reliability	Process in development	Ongoing reliability process, but not implemented because of high assessor turnover

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

ECE = early care and education

Key revisions from 2011 to 2014

- Refined and clarified standards based on the pilot and stakeholders' feedback. For example, the state modified the developmentally appropriate practice component to focus on practices at the classroom rather than program level.
- Adjusted the rating level at which certain standards applied. For example, programs were required to have individualized activities to meet the developmental needs of all children informed by an IFSP/IEP as applicable at level 1; previously this standard was applied only at higher ratings levels.
- Modified use of ERS and CLASS scores to target program improvement efforts rather than to determine rating levels. Findings from the pilot revealed that it was challenging for programs to meet the specified cut scores for ERS and CLASS at levels 4 and 5.
- Initiated monitoring visits to verify the documentation used to meet the standards in rated programs. The visits also served to identify areas for ongoing improvement and technical assistance needs.
- Began work to align the standards of public prekindergarten programs with the revised Maryland EXCELS standards for child care centers.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of center-based early care and education programs were provided by the state.

NOTES:

[a] We use the term "level" across states, rather than the state-specific term used for rating levels.

[b] In order to use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout el. al, 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Massachusetts Quality Rating and Improvement System

Grant details

Grantee: Massachusetts Department of Early Education and Care (EEC)

TQRIS lead agency: EEC

RTT-ELC grant amount: $50,000,000

TQRIS implementation status

Pilot: 2010–2011

Implementation date: 2011 (statewide implementation)

Date of current TQRIS standards: January 2014

TQRIS structure[a]

Defining rating levels: Building-block system with four levels

Entry into and progression through TQRIS:

- Licensing compliance is the only requirement at level 1. Public schools and license-exempt programs are eligible and do not have to become officially licensed to participate.
- Programs may apply to enter and advance at any level; however, all programs are encouraged, but not required, to enter at level 1 and progress through each rating.
- No minimum or maximum time requirement is imposed at any rating level.
- Ratings are valid for two years.

TQRIS participation

Participation requirements: Enrollment is mandatory for Head Start programs, public school and community based-programs receiving Universal Prekindergarten (UPK) and Massachusetts Inclusive Preschool Learning Environments grant funds, and programs receiving Child Care and Development Fund (CCDF) subsidies; participation is voluntary for all other programs.

Participation of early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
Programs receiving CCDF subsidies	3,702 (83%)	n.a.
Head Start and Early Head Start programs	233 (100%)	n.a.
UPK programs	224 (100%)	
Programs funded by IDEA, Part B, section 619	103 (20%)	n.a.
Programs funded under Title I	20 (10%)	n.a.
Programs funded by Massachusetts Inclusive Preschool Learning Environments grants	136 (100%)	n.a.
License-exempt programs	122 (52%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include both center-based and family child care programs.

IDEA = Individuals with Disabilities Act; CCDF = Child Care and Development Fund; UPK = Universal Prekindergarten; n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of ratings)	Number of children (percent of children in rated programs)
Level 1	4,589 (74%)	n.a.
Level 2	1,497 (24%)	n.a.
Level 3	94 (2%)	n.a.
Level 4	2 (<1%)	n.a.

Note: Counts include both center-based and family child care programs. The state began granting level 3 and 4 ratings in January and November 2014, respectively. As reported in Massachusetts' RTT-ELC 2014 Annual Performance Report, some programs receive multiple ratings because their applications were verified at each rating level. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Massachusetts reports that the actual number of programs participating in the TQRIS is 5,891 (including an unspecified number of programs that have applied but have not yet been rated). Percentages are calculated based on total ratings (6,182).

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- Head Start and programs accredited by the National Association for the Education of Young Children (NAEYC) receive automatic credit for several rating criteria.

Incentives:

- Programs at level 2 or above serving infants and toddlers are eligible for tiered reimbursement rates of 3 percent above the standard daily rate.
- The state offers QRIS Improvement Grants and Child Care Quality Grants that may be used for program planning and durable goods; grants range from $4,500 to $10,500 based on enrollment. Child Care Quality grants may be used to fund one-time expenses such as accreditation fees, training, or materials.
- To be eligible for Universal Prekindergarten funding, Head Start state supplemental funding, and Massachusetts Inclusive Preschool Learning Environment grants, programs must be enrolled in TQRIS.

Quality components[b]

Component categories: Licensing compliance, health promotion and safety, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, administration and management, community involvement, cultural and linguistic diversity, provisions for special needs.

Observational measures: Environment Rating Scale (ERS) observations are conducted. Programs must also complete self-assessments by using the (1) Program Administration Scale, (2) Strengthening Families Self-Assessment, and (3) Arnett Caregiver Interaction Scale or Classroom Assessment Scoring System.

Generating ratings

Data collection, verification, and calculation of ratings: Programs must submit documentation in support of all requirements except for licensing. Licensing compliance is verified through the state licensing database. Raters review and verify evidence pertaining to other requirements within TQRIS, and the system keeps a tally of the standards met. When all criteria are verified, TQRIS automatically calculates a rating. Level 4 requires a classroom observation by a state-certified observer. Programs may request up to four exemptions (valid for one year) for standards that prevent them from achieving a higher rating. Programs must develop a plan and timeline to meet the standards for which they are requesting an exemption.

Staff qualifications, training, and reliability requirements

	Program raters	Classroom observation assessors
Type of staff	EEC staff	Contracted staff
Required education/ qualifications	A bachelor's degree with a major in education or special education, educational administration, or educational counseling OR Two years of full-time, or equivalent part-time, professional experience in teaching, educational administration, counseling, or guidance	No state-specified education requirements; assessor must be a certified ERS reliable rater trained by the authors within the past five years; demonstrate experience in administering ERS tools as well as experience with the state's early care and education delivery system; and be knowledgeable of EEC's licensing regulations.
Training/support	Six-month training process, including shadowing a veteran rater; documentation verification manual guides rating process.	Trained by state anchor or tool developer*
Initial reliability requirement	No formal process is in place. TQRIS director reviews all level 4 ratings.	90 percent against anchor over three visits
Ongoing reliability requirement	No formal process is in place.	90 percent against anchor; assessed quarterly

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

EEC = Early Education and Care; ERS = Environment Rating Scales

Key revisions from 2011 to 2014

- Required enrollment in the TQRIS for programs that receive Child Care and Development Fund subsidy payments.
- Revised select standards such as requiring that all professional development activities used to meet the TQRIS training criteria must also be eligible for continuing education units or college credit.
- No longer required an ERS observation at level 3 and only used an ERS observation in the rating determination at level 4. This shifted the focus of the observation at level 3 to program development; programs completed a self-assessment at level 3 and program quality specialists conducted technical assistance visits to help prepare programs for the external ERS observation to achieve level 4.
- Began using the ERS developer's alternative method for scoring in which all indicators within a section are scored to get an overall score, rather than assigning a low rating for an entire section if it rated low on a single indicator. This change minimized the effect of some indicators, such as not having a separate bathroom for staff, which previously depressed subscale scores.
- Began implementing additional modifications to the standards. For example, the requirement for programs to have an annual health consultant visit will be moved from level 2 to level 3 due to a lack of health consultant capacity and to provide assistance to programs at level 2 to prepare for the visits at level 3.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Minnesota Parent Aware

Grant details

Grantee: Minnesota Office of Early Learning

TQRIS lead agency: Minnesota Department of Human Services

RTT-ELC grant amount: $44,858,313

TQRIS implementation status

Pilot: 2007–2011

Implementation date: January 2015 (statewide implementation)

Date of current TQRIS standards: 2012

TQRIS structure[a]

Defining rating levels: Hybrid system with four levels.

Entry into and progression through TQRIS:

- Licensing compliance is a prerequisite for all programs except for those using an accelerated pathway.
- Programs may apply for a specific rating level (at entry or to progress).
- The state uses a cohort system to manage participation in the full rating process. Once a program receives a rating, it must wait six months until the next twice yearly cohort entry date to apply for a higher rating. Ratings are published six months after application; thus, it takes at least one year to receive a new rating. There is no maximum for how long a program may remain at any level.
- Ratings are valid for two years.

TQRIS participation

Participation requirements: Participation is voluntary for all programs; license-exempt programs are not eligible.

Participation of early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
School-based preschool programs	548 (89%)	n.a.
Head Start and Early Head Start programs	257 (90%)	n.a.
Programs funded by IDEA, Part C	8 (7%)	n.a.
Programs funded by IDEA, Part B, section 619	29 (9%)	n.a.
Programs funded under Title I	158 (92%)	n.a.
Programs receiving CCDF subsidies	531 (17%)	n.a.
Licensed centers and family child care program not receiving CCDF subsidies	660 (8%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include both center-based and family child care programs.

IDEA = Individuals with Disabilities Education Act; CCDF = Child Care and Development Fund; n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	227 (12%)	n.a.
Level 2	268 (14%)	n.a.
Level 3	72 (4%)	n.a.
Level 4	1,325 (70%)	n.a.

Note: Counts include both center-based and family child care programs.

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- An accelerated pathway to rating is available to Head Start, public prekindergarten, and nationally accredited programs. To earn a level 4 rating, these programs must use an approved curriculum, use an approved child assessment tool, and deliver training to lead teachers in implementing curriculum and child assessments. Programs may apply for ratings at any time and are not subject to the twice yearly cohort entry process.

Incentives:

- Level 3- and 4-rated programs are eligible for tiered child care subsidy reimbursement.
- Prerating support, up to $500, is available to licensed, nonaccredited programs, and up to $1,000 is available for achievement of a level 1, 2, or 3 rating.
- Early Learning Scholarships are available to families with high needs to promote access to Parent Aware participating programs. Programs receive $3,000 to $5,000 per child per year based on rating level.

Quality components[b]

Component categories: Licensing compliance, health promotion and safety, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, community involvement, cultural and linguistic diversity, provisions for special needs.

Observational measures: The Pre-K CLASS is used in classrooms where at least 50 percent of children are age 36 months or older. Classrooms serving younger children are not observed. Public prekindergarten, Head Start, and accredited programs are also not observed.

Generating ratings

Data collection, verification, and calculation of ratings: Programs submit documentation to verify all standards except for the observation requirement. Classroom observations are required for licensed, nonaccredited programs seeking a level 3 or 4 rating. Within the data system, raters record whether a standard was met and then award points. The system automatically calculates the final rating.

Staff qualifications, training, and reliability requirements

	Program raters	Observational measure assessors
Type of staff	State-certified staff from the Minnesota Department of Education (DOE) process ratings for Head Start, public prekindergarten, early childhood special education, and charter school programs. Contracted raters process ratings for licensed center-based programs.	Contracted staff
Required education/ qualifications	No state-specified education requirements.	No state-specified education requirements.
Training/support	A test version of the TQRIS data system helps raters become familiar with the process and includes practice rating programs. Raters refer to a flowchart for each standard to guide the rating process.	Two-day training by anchor contracted staff or CLASS developer. Ongoing training including quarterly group scoring meetings and participation in professional development opportunities.
Initial reliability	Exact agreement with anchor rater on 85 percent of the standards across five ratings*	Assessors must pass a web-based CLASS reliability test and achieve a live observation reliability score of 85 percent or above.
Ongoing reliability	The anchor rater assesses at least 10 percent of ratings, and exact agreement on 85 percent of all standards must be achieved.	Ongoing reliability is assessed quarterly or every 10 visits (whichever comes first), and assessors must achieve 85 percent agreement within one point. Assessors must also recertify annually as a CLASS observer and trainer.

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

CLASS = Classroom Assessment Scoring System

Key revisions from 2011 to 2014

- Changed from a points to a hybrid rating structure to increase variation in the ratings based on findings from the pilot that showed a skewed rating distribution toward the higher rating levels.
- Revised the content and number of standards based on findings from the pilot. For example, the state dropped a family partnership component due to lack of variation in scores and allowed coaching and mentoring to count toward training requirements.
- Switched from the ERS to the CLASS observation measure for use in rating determination at levels 3 and 4.
- Moved from automatically granting Head Start, public prekindergarten, and accredited programs a level 4 rating to an accelerated pathway for these programs. The accelerated pathway waived TQRIS standards that align with existing standards in place for the programs that were eligible for this pathway (such as Head Start Program Performance Standards or accreditation standards). Programs that participated in the alternative pathway must still meet TQRIS standards related to curriculum and the use of child assessments.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

North Carolina Star Rated License System

Grant details

Grantee: Governor's Office, Early Childhood Advisory Council

TQRIS lead agency: North Carolina Division of Child Development and Early Education (DCDEE)

RTT-ELC grant amount: $69,991,121

TQRIS implementation status

Pilot: No pilot

Implementation date: 1999 (statewide implementation)

Date of current TQRIS standards: January 2013

TQRIS structure[a]

Defining rating levels: Point system with five levels

Entry into and progression through TQRIS:

- All licensed programs automatically receive a level 1 rating. Licensing compliance is the only requirement at level 1. Programs may then choose to apply for a higher rating level.
- Level 1 programs must have a valid license for six months before they may apply for a higher rating level. There is no limit to how long a program may remain at any level.
- Ratings are valid for three years.

TQRIS participation

Participation requirements: Mandatory at level 1 for all licensed programs (Head Start and school-sponsored prekindergarten programs must be licensed); voluntary for nonlicensed programs such as federally funded school-based prekindergarten programs (these types of programs must become licensed to participate).

Participation of early care and education (ECE) programs (as of December 2014)

Type of Program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
State-funded prekindergarten programs (includes school- and community-based programs)	1,170 (100%)	n.a.
Head Start and Early Head Start programs	402 (95%)	n.a.
Programs funded by IDEA, Part B, section 619	1,308 (65%)	n.a.
Programs funded under Title I	510 (79%)	n.a.
Programs receiving CCDF subsidies	4,952 (95%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include both center-based and family child care programs.

IDEA = Individuals with Disabilities Education Act; CCDF = Child Care and Development Fund; n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	441 (6%)	n.a.
Level 2	270 (4%)	n.a.
Level 3	1,630 (22%)	n.a.
Level 4	1,803 (25%)	n.a.
Level 5	2,302 (44%)	n.a.

Note: Counts include both center-based and family child care programs.

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- There are no alternative pathways.

Incentives:

- Licensed programs with a level 3 rating or higher are eligible to receive child care subsidy payments. Tiered reimbursement rates are available to the same programs.
- Achievement of a level 4 rating or higher is an eligibility requirement for NC Pre-K, the state-funded prekindergarten program.
- Some local partnerships of Smart Start, a statewide public/private partnership to improve North Carolina's system of early care and education, offer support such as bonuses for achieving a higher rating or funds for materials, but the level of available support varies across the state.

Quality components[b]

Component categories: Licensing compliance, health promotion and safety, ratio and group size, curriculum, environment, workforce qualifications and credentials, family engagement, administration and management.

Observational measures: An Environment Rating Scales (ERS) assessment is required to earn three points in the program standards domain (13 total points are required to achieve a level 5 rating).

Generating ratings

Data collection, verification, and calculation of ratings: To verify achievement of program standards, licensing consultants review programs' paper applications and conduct site visits, which involve reviewing files, observing classrooms, and monitoring staff-child ratios and group sizes. To verify achievement of education standards, the licensing consultant confirms that staff's requisite credentials appear in the early childhood workforce (ECW) database. Contracted staff conduct ERS observations. The rater calculates the total number of points earned and provides data entry staff with the rating information for entry into the database. The data system also generates a total score, and the data entry staff check to ensure that the total scores match.

Staff qualifications, training, and reliability requirements

	Program raters	Classroom observation assessors
Type of staff	DCDEE staff	Contracted staff
Required education/ qualifications	Bachelor of arts (BA) required, preferably in early care and education; three years of experience as a center administrator.	BA required, preferably in early care and education; a minimum of one year of teaching experience in an early childhood program.
Training/support	New raters complete sample ratings of various program types and shadow a current rater. Raters follow a procedures manual.	The contractor trains the assessors.
Initial reliability requirement	There is no formal process. The supervisor tracks the rater's accuracy and error percentages over time and provides additional training as needed.	Average of 85 percent across three observations measured against state anchor, with the last two scores at least 85 percent.*
Ongoing reliability requirement	There is no formal process.	The first two ongoing reliability checks occur on or before the 6th assessment. Then, the average across the last three checks determines the frequency of ongoing checks (with an 85 percent minimum required).

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

DCDEE = Division of Child Development and Early Education; BA = bachelor of arts

Key revisions from 2011 to 2014

- Refined standards to require that level 4 and 5 programs serving 4-year-old children use an approved curriculum and to provide options for ways programs could meet the lead teacher education requirements.
- Required programs that receive Child Care and Development Fund subsidy payments to achieve a level 3 rating or higher.
- NC Pre-K, the state-funded prekindergarten program, required programs to achieve a level 4 rating or higher to receive NC Pre-K funding.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Ohio Step Up to Quality Program

Grant details

Grantee: Office of the Governor, State of Ohio Office of Early Learning and School Readiness

TQRIS lead agency: Ohio Department of Job and Family Services (JFS) and Department of Education (DOE)

RTT-ELC grant amount: $69,993,362

TQRIS implementation status

Pilot: 2004 - 2005

Implementation date: 2006 (statewide implementation)

Date of current TQRIS standards: October 2013

TQRIS structure[a]

Defining rating levels: Hybrid system with five levels; first three levels are blocks, and levels 4 and 5 are points.

Entry into and progression through TQRIS:

- All programs must be licensed to achieve level 1.
- All programs may apply for a specific rating level (at entry or to progress).
- Programs must wait a minimum of six months to apply for the next rating level. Programs are not subject to a maximum time limit for remaining at a particular level.
- Level 1 ratings are valid for one year; level 2 and 3 ratings are valid for two years; level 4 and 5 ratings are valid for three years. All programs must submit an annual report on continuous quality improvement to maintain their ratings.

TQRIS participation

Participation requirements: Participation is mandatory for all school-operated early childhood programs and programs with children receiving Child Care Development Fund (CCDF) subsidies. Participation is mandatory for all programs licensed through DOE and receiving state funding for chartered nonpublic and public preschools. Participation is voluntary for programs licensed through JFS unless they receive Early Childhood Education Expansion grant funds.

Participation of early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
State-funded preschool	159 (41%)	n.a.
Early Head Start and Head Start programs	283 (56%)	n.a.
Programs funded by IDEA, part B, section 619	119 (24%)	n.a.
Programs receiving CCDF subsidies	1,029 (35%)	n.a.
Licensed programs not receiving CCDF subsidies	391 (24%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include both center-based and family child care programs.

IDEA = Individuals with Disabilities Education Act; CCDF = Child Care Development Fund; n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	639 (39%)	n.a.
Level 2	130 (8%)	n.a.
Level 3	380 (23%)	n.a.
Level 4	279 (17%)	n.a.
Level 5	202 (12%)	n.a.

Note: Counts include both center-based and family child care programs.

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- No alternative pathway.

Incentives:

- Ohio offers tiered reimbursement rates that increase with the rating level and the proportion of subsidized children served.
- Ohio offers a Quality Achievement Award to programs once they achieve a particular rating level. The amount varies by rating level, with higher-level programs receiving greater amounts.

Quality components[b]

Component categories: Licensing compliance, ratio and group size, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, administration and management, accreditation, community involvement.

Observational measures: Ohio Classroom Observation Tool (OCOT); required only for levels 3 through 5.

Generating ratings

Data collection, verification, and calculation of ratings: State raters complete a screening of programs' licensing status, followed by a desk review of documentation submitted by programs. Raters then conduct an unannounced on-site verification visit to observe classrooms, review documentation, and interview lead teachers and administrators. Raters submit recommended ratings to state supervisors and, once approved, email the rating to programs. State-certified raters calculate the final ratings.

Staff qualifications, training, and reliability requirements

	Program raters	Observational measure assessors
Type of staff	Staff from JFS and DOE conduct OCOT observations and rate programs.	
Required education/ qualifications	The Department of Administrative Services (human resources for all state agencies) sets the JFS qualifications. If an applicant has no bachelor of arts, he or she may be hired based on experience. DOE requires a master's degree in early education or a related field. JFS and DOE qualifications include experience as a program administrator/director, experience in early childhood, and experience in using the observation tool.	
Training/support	Staff are trained according to a standardized process, and raters shadow current raters,	Master trainers from JFS and DOE train staff. Staff participate in a two-day training and have the opportunity to complete sample observations in a variety of settings.
Initial reliability requirement	85 percent agreement on each element and 100 percent agreement on overall rating with anchor for three consecutive visits*	85 percent agreement on each element and 100 percent agreement on overall rating with anchor for three consecutive visits*
Ongoing reliability requirement	85 percent agreement on each element and 100 percent agreement on overall rating with anchor across all subsequent visits State anchors conduct random reliability checks of reliable observers twice annually. Two raters from both agencies complete visits together and review results.	85 percent agreement on each element and 100 percent agreement on overall rating with anchor across all subsequent visits State anchors conduct random reliability checks of reliable observers twice annually.

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

JFS = Job and Family Services; DOE = Department of Education; OCOT = Ohio Classroom Observation Tool

Key revisions from 2011 to 2014

- Transitioned from a three-level building block rating structure to a five-level hybrid structure and significantly revised the TQRIS standards, including adding a family and community partnerships component.
- Developed and implemented a classroom observation measure for use in the rating determination.
- Expanded TQRIS eligibility to include DOE public prekindergarten programs.
- Instituted review of program-submitted documents before conducting classroom observations. This was intended to enhance the efficiency of the verification process by first determining whether programs meet the minimum requirements for the rating level at which observations are required.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Rhode Island BrightStars

Grant details

Grantee: Rhode Island Department of Education (RIDE)

TQRIS lead agency: Rhode Island Department of Human Services

RTT-ELC grant amount: $50,000,000

TQRIS implementation status

Pilot: 2007–2008

Implementation date: 2009 (statewide implementation)

Date of current TQRIS standards: September 2013

TQRIS structure[a]

Defining rating levels: Building-block system with five levels.

Entry into and progression through TQRIS:

- Licensing compliance plus development of a quality improvement plan is required at level 1. School-sponsored preschool programs are eligible to participate if they meet RIDE compliance requirements.
- All programs may apply for any specific rating level (at entry or to progress).
- Programs may remain at each rating level for any length of time (no minimum or maximum time requirement at any rating level).
- Ratings are valid for three years, but programs must update their program improvement plan annually.

TQRIS participation

Participation requirements: Enrollment is mandatory for programs receiving Child Care and Development Fund subsidies. Participation is voluntary for all other programs.

Participation of center-based early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
School-sponsored preschool programs*	15 (28%)	1,111 (3%)
State-funded preschool programs located in community-based centers	11 (92%)	268 (1%)
Head Start programs	36 (90%)	1,994 (6%)
Licensed, center-based programs	245 (79%)	8,731 (25%)
Accredited, center-based programs	24 (86%)	1,186 (3%)

*School-sponsored preschool programs became eligible to participate when the revised standards went into effect in fall 2013.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating.

Distribution of programs and children across rating levels (as of December 2014)

	Number of center-based ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	113 (34%)	3,121 (23%)
Level 2	63 (19%)	2,234 (17%)
Level 3	48 (15%)	2,338 (18%)
Level 4	68 (21%)	3,768 (28%)
Level 5	39 (12%)	1,820 (14%)

Note: Program counts by level were calculated based on state-reported data on the number of programs in each level by program type. Because program type categories were determined by funding source or operating authority, programs with more than one funding stream and/or operating authority may be counted twice at a particular level.

ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- Head Start and nationally accredited programs receive an automatic top rating on 3 of the 10 standards in curriculum, assessment, and family communication and involvement.

Incentives:

- BrightStars offers an annual quality award for programs with a level 3, 4, or 5 rating. Awards range from $50 to $500 per child per year based on overall enrollment, the number of children receiving subsidized care, and the program's rating level.
- Programs may apply for quality improvement grants to address goals identified in their improvement plans.
- T.E.A.C.H. scholarships are only available to staff in programs participating in BrightStars.

Quality components[b]

Component categories: Licensing compliance, ratio and group size, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, provisions for special needs.

Observational measures: Environmental Rating Scales (ERS) and Pre-K Classroom Assessment Scoring System (CLASS) assessments are conducted. All ERS subscales are assessed except for the parent-related items from the parents and staff subscale. A Pre-K CLASS observation must be conducted, but programs do not have to achieve a specific score.

Generating ratings

Data collection, verification, and calculation of ratings: Programs complete a paper application and attach supporting documentation. (A new data system is in development; it will allow online application.) At level 2, program raters conduct site visits to observe ratios and group sizes and to review documentation such as professional development plans. ERS classroom observations are required beginning at level 3, and a CLASS observation is conducted at level 5.

Staff qualifications, training, and reliability requirements

	Program raters	Classroom observation assessors
Type of staff	Contracted staff (the same organization is responsible for the ratings and observations).	Contracted staff (the same organization is responsible for the ratings and observations).
Required education/ qualifications	Bachelor of arts required, preferably in early care and education; experience in data collection and program evaluation and professional experience in the education field (not specified).	BA required, preferably in early care and education
Training/support	Raters conduct mock document reviews and shadow a current rater. A policy and procedures manual guides the rating process.	Trained by ERS authors unless time constraints require in-house training; CLASS assessors attend the state CLASS training if not already certified.
Initial reliability requirement	No formal process. The assessment manager reviews all ratings completed by new raters and conducts random rating checks.	ERS: 85 percent against ERS authors across three observations; CLASS: no reliability policy.
Ongoing reliability requirement	No formal process.	ERS: 85 percent against a reliable assessor across last three reliability checks; checked every 20th observation for author-trained assessors, every 10th observation for in-house trained assessors

ERS = Environment Rating Scales; CLASS = Classroom Assessment Scoring System

Key revisions from 2011 to 2014

- Created a new set of TQRIS standards aligned with licensing and public prekindergarten standards.
- Revised certain standards that seemed to prevent programs from achieving high rating levels as a result of simulations that examined the distribution of ratings among programs if the standards were changed. For example to achieve a level 5, 75 percent of teachers must meet the specified educational requirements rather than "all teachers" as was previously required.
- Adjusted the rating level at which certain standards applied, such as requiring the use of a developmentally appropriate curriculum to achieve a level 3 rating or higher, when it was previously required only for level 4 and higher.
- Expanded TQRIS eligibility to include public school prekindergarten programs and required programs receiving Child Care and Development Fund subsidies to enroll in BrightStars.
- Updated the documentation requirements for specific standards, such as requiring submission of programmatic curriculum outlines instead of classroom-level curriculum outlines.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015.Data on the distribution of programs and children across rating levels and participation of center-based early care and education programs was provided by the state.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

Washington Early Achievers

Grant details

Grantee: Washington State Department of Early Learning

TQRIS lead agency: Washington State Department of Early Learning

RTT-ELC grant amount: $60,000,000

TQRIS implementation status

Pilot: 2008–2011 and 2012–2013

Implementation date: 2012 (statewide implementation)

Date of current TQRIS standards: June 2012

TQRIS structure[a]

Defining rating levels: Hybrid system with five levels; blocks for levels 1 and 2 and points for levels 3 through 5.

Entry into and progression through TQRIS:

- Licensing or certification compliance is the only requirement at level 1.
- Except for programs eligible for the alternative pathways all programs enter at level 1; to achieve level 2, programs must complete a series of activities and then may apply for a specific higher rating level.
- Programs may remain at each rating level for any length of time (no minimum or maximum time at any level).
- Ratings are valid for three years.

TQRIS participation

Participation requirements: Mandatory for Early Childhood Education and Assistance Program (ECEAP) programs, the state-funded prekindergarten program for children from low-income households; voluntary for all other programs.

Participation of early care and education (ECE) programs (as of December 2014)

Type of program	Number of participating programs (percent of eligible programs)	Number of children served (percent of preschool-age children in state)
ECEAP programs	260 (100%)	n.a.
Early Head Start and Head Start programs	415 (100%)	n.a.
Licensed, center-based programs	1,567 (100%)	n.a.
Programs receiving CCDF subsidies	4,718 (100%)	n.a.

Note: Programs with more than one funding stream and/or operating authority were reported within each relevant category. As a result, some programs may be counted in multiple rows and the total across categories may be larger than the actual number of programs participating. Counts include both center-based and family child care programs. As reported in Washington's RTT-ELC 2014 Annual Performance Report, all programs that are licensed or are a Head Start/ECEAP site are counted as participating in TQRIS.

ECEAP = Early Childhood Education and Assistance Program; CCDF = Child Care and Development Fund; n.a. = not available

Distribution of programs and children across rating levels (as of December 2014)

	Number of ECE programs (percent of rated programs)	Number of children (percent of children in rated programs)
Level 1	3,367 (58%)	n.a
Level 2	2,004 (34%)	47,936 (n.a.)
Level 3	351 (6%)	21,065 (n.a.)
Level 4	92 (2%)	7,354 (n.a.)
Level 5	1 (<1%)	268 (n.a.)

Note: Counts include both center-based and family child care programs.

n.a. = not available; ECE = early care and education

Incentives to promote participation

Accelerated or alternative pathway to ratings:

- Washington offers alternative pathways for ECEAP and Head Start programs. If 75 percent or more of total facility slots are ECEAP or Head Start slots, then the program uses the alternative pathway. Programs enter at level 3, but they do not receive an official rating until they are observed. Programs may earn a level 2, 4, or 5 rating.

Incentives:

- The state offers tiered reimbursement rates, with higher-rated programs receiving higher rates of reimbursement.
- Level 3 through 5 programs receive quality improvement awards, which vary by level ($5,000 to $9,000).
- Level 2 programs serving low-income children receive up to $500 for materials and other supports to help them with their Environment Rating Scale (ERS) assessments.
- The state awards professional development scholarships; employees at centers with a level 3 rating or higher are eligible.

Quality components[b]

Component categories: Licensing compliance, curriculum, child assessment, environment, workforce qualifications and credentials, family engagement, administration and management.

Observational measures: A facility must achieve minimum specified ERS and CLASS scores in order to achieve a rating of level 3. ERS and CLASS scores make up 55 percent of the total possible points, with 40 percent determined by CLASS scores and 15 percent by ERS scores.

Generating ratings

Data collection, verification, and calculation of ratings: Programs must first register and submit an application within the state database. After completing level 2 activities, programs register for an on-site evaluation period window. An unannounced on-site evaluation occurs within the scheduled window and includes classroom observations, staff and parent interviews, and records and documentation review. After the observation component is complete, two raters code program records for points. If the two raters reach agreement, the program is awarded a final rating. Any disagreements are subject to discussion among both coders and a lead staff.

Staff qualifications, training, and reliability requirements

	Program raters	Observational measure assessors
Type of staff	Contracted staff; program raters and assessors are the same staff.	
Required education/ qualifications	Bachelor of arts (BA) required; early care and education, preferred. Exceptions in education requirements are made for those with language skills (for example, BA requirement waived if rater can speak Spanish). Qualifications include two years of experience in early education and experience in large-scale research projects. Raters/assessors who speak several languages are preferred.	
Training/support	Records review training includes overview of criteria, practice with example files, and practice with second-coding documents. Rater also shadow current raters.	In-person training by the tool's author and an official tool anchor and completion of online training through the tool publisher.*
Initial reliability requirement	All records are double-coded, so there is no set level of reliability.	ERS assessors must achieve 85 percent agreement with anchor across three reliability checks; CLASS assessors must achieve 80 percent agreement during on online training.
Ongoing reliability requirement	All records are double coded, so there is no set level of reliability.	ERS assessors: 90 percent with anchors every 11th assessment or once per month. CLASS assessors: 80 percent online training repeated annually.

*Anchors are observers who have shown consistent reliability on the observation measure and serve as the reliability benchmark for other assessors.

BA = bachelor of arts; ERS = Environment Rating Scales; CLASS = Classroom Assessment Scoring System

Key revisions from 2011 to 2014

- Adjusted standards to increase the focus on those related to family engagement and individualization of services.
- Revised the scoring process for observational measures to require that an average of all classroom scores on the CLASS and ERS within a program meet required thresholds, rather than requiring a minimum score for each classroom.
- Extended TQRIS eligibility to and implemented an accelerated pathway for Head Start and ECEAP programs.
- Expanded the coaching structure by offering coaching to level 2 programs in addition to level 3 and 4 programs.

Source: Document reviews and telephone interviews conducted by Mathematica, October 2014–April 2015. Data on the distribution of programs and children across rating levels and participation of early care and education programs are from the state's RTT-ELC 2014 Annual Performance Report.

NOTES:

[a] We use the term "level" across states rather than the state-specific term for rating levels.

[b] To use consistent terms across states, we use the 13 quality components defined by the *Compendium of Quality Rating Systems and Evaluations* (Tout et al. 2010) rather than state-specific categories.

This page has been left blank for double-sided copying.

This page has been left blank for double-sided copying.


DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA