

Quality Rating and Improvement Systems for Early Care and Education Programs: Making the Second Generation Better

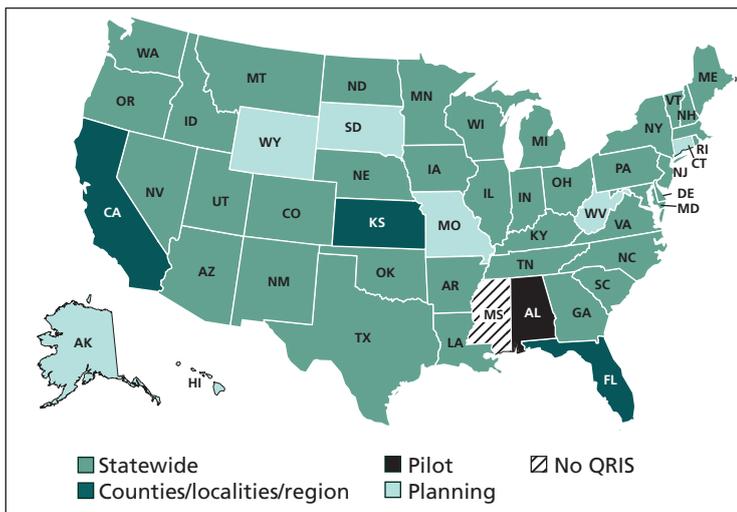
Jill S. Cannon, Gail L. Zellman, Lynn A. Karoly, and Heather L. Schwartz

An extensive body of research in child development, neuroscience, and other disciplines demonstrates the significance of the period from birth to school entry for children's development and calls attention to the importance of the quality of the early care and education (ECE) experiences young children receive.¹ Although each state in the United States has a licensure system intended to ensure that ECE programs provide safe and secure settings, these requirements have historically imposed a fairly low threshold for the quality of the care and education environment, focused as they are on physical safety.² The neuroscience and child development research, in conjunction with findings in the 1990s that most nonparental care was mediocre in quality at best, led decisionmakers to search for ways to improve the quality of care in ECE programs.³ Quality rating and improvement systems (QRISs), which treat quality of care in a multidimensional way, began at the end of the 1990s and have now been

almost universally adopted as one tool that states and localities have employed to boost quality in ECE programs. Today, 49 states have a QRIS—either implemented, piloted, or planned in some or all parts of the state—thanks in part to federal incentives (Figure 1).

QRISs assess quality based on multiple indicators, such as group size, staff-child ratio, teacher and administrator qualifications, and elements of the care and education environment. These quality indicators are then combined to produce a simple quality rating metric, such as a one- to five-star rating. To motivate and help ECE programs move up in the rating system, a QRIS generally offers ECE programs improvement supports: technical assistance, professional development, or child care subsidy reimbursements tied to the program's rating level. The basic QRIS approach is to establish quality standards at successively higher levels, measure the quality level that programs reach given those standards, provide supports and financial incentives for programs

Figure 1. Status of QRISs Across the United States



SOURCE: QRIS National Learning Network, "QRIS State Contacts and Map," undated, last updated January 2017.

to achieve and sustain higher levels of quality, and disseminate objective information about quality that parents can use in selecting an ECE program.

QRISs are at a critical point in their development and implementation. A wave of QRIS evaluations, most of which are validation studies (discussed in more detail later in this perspective), is becoming available, largely funded through three rounds of federal Race to the Top–Early Learning Challenge (RTT–ELC) grants covering QRISs in 20 states.⁴ The RTT–ELC grants have also been a key source of funds for the development and expansion of ECE QRISs. As the federal grant funds expire, states will have fewer resources available to operate their QRISs without new sources of funding. States will need to be more strategic about the alloca-

tion of funds for and within these systems to achieve their goals of expanding access to and improving the quality of ECE programs.

In this perspective, we suggest some ways to accomplish this. We assess what the early childhood field has learned about QRISs as they have been widely adopted and matured, and how the field can strategically move this first generation of QRISs into a second generation. Although QRISs may include before- and after-school programs for school-age children, we focus on QRISs as they are applied to child care and early education programs serving children from birth to kindergarten entry.

Background on QRIS Development

QRISs are essentially ECE accountability systems centered on quality ratings. In a given QRIS, a summary quality rating for each participating program is derived by combining measures of program performance on a set of elements that are believed to jointly define quality. These elements are typically assessed with a range of measures evaluating structural characteristics, such as group size, staff-child ratios, and teacher qualifications, and measures of the caregiving process, such as assessments of teacher responsiveness to children's needs.

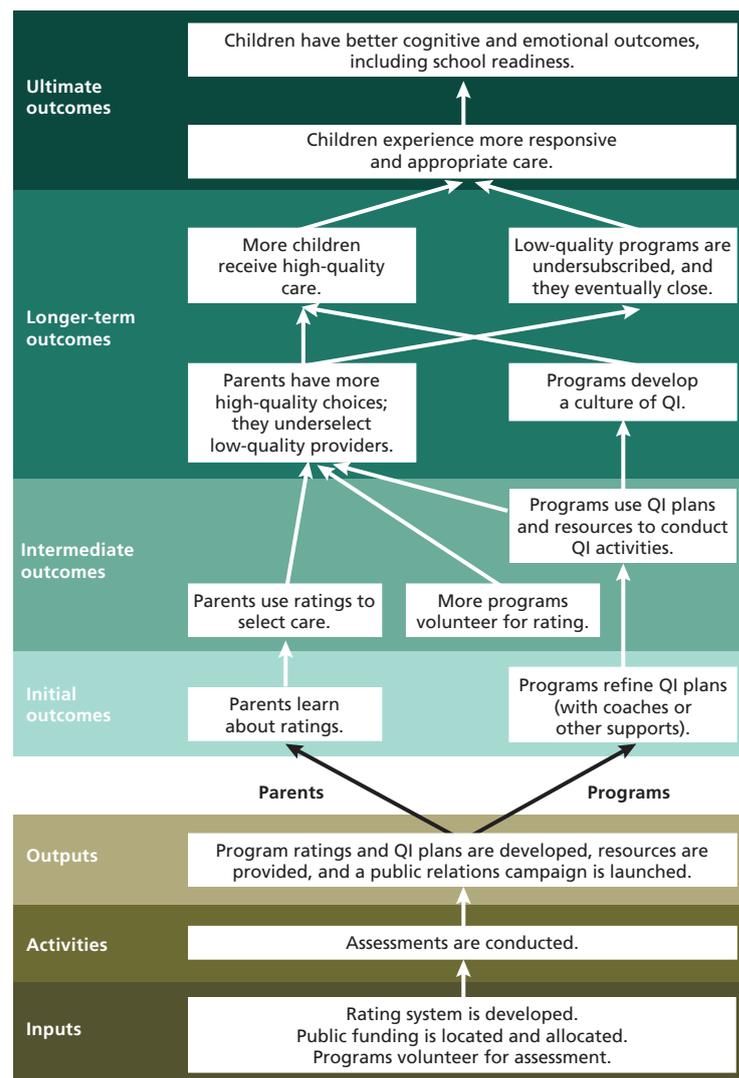
The genesis of QRISs was a concern that ECE quality was generally inadequate in terms of providing warm and responsive care in safe, well-managed, and developmentally appropriate environments; that licensing standards were, by design, not addressing quality of care; and that public funds, usually through subsidies (i.e., the federal Child Care and Development Fund block grant program), were supporting poor-quality ECE programs. The rating process and its outputs were designed to improve child care quality by defining quality standards; making program qual-

ity transparent to consumers, providers, and policymakers; and providing incentives and supports for quality improvement toward the ultimate goal of improving children’s outcomes. Early QRIS developers believed that the QRIS approach could and should be used in a range of ECE settings: both center-based and home-based care, for community-based and school-based child care and early education programs, and for children from birth through school age. Although the recent rounds of federal grants for ECE have encouraged statewide systems, some QRISs have been designed and implemented at a local level, typically by counties.

Rating systems are premised on a market model, where program ratings are expected to incentivize ECE providers to improve the quality of the ECE experience that they deliver. An underlying assumption of the QRIS approach is that parents have choices in selecting an ECE program for their child, but they may have difficulty judging program quality. With the information from the QRIS ratings, parents have a tool to judge quality and are expected to choose a higher-rated program among those that they can afford and that meet other needs they have for care (e.g., accessibility, operating hours). As parents select better-rated programs subject to their budget constraints, programs with lower ratings will be motivated to improve. Figure 2 provides an example of the general logic model for a rating system.

As QRISs have evolved, their goals have expanded. While QRISs were originally designed primarily to inform parental choices and thereby motivate programs to improve quality, decisionmakers have come to see that QRISs must also include quality improvement supports and additional incentives to improve and sustain quality, especially in communities where parents cannot afford to pay the full cost of higher-quality programs. Systems

Figure 2. QRIS Logic Model



SOURCE: Gail L. Zellman and Michal Perlman, *Child-Care Quality Rating and Improvement Systems in Five Pioneer States*, Santa Monica, Calif.: RAND Corporation, MG-795-AECF/SPF/UWA, 2008. NOTE: QI = quality improvement.

are also increasingly embracing other objectives, such as professionalizing the ECE workforce (e.g., through well-specified career pathways) or ensuring that providers are commercially viable (e.g., through rating standards focused on following good business practices). These new objectives are understandable efforts to improve the overall landscape for the provision of high-quality care, but they threaten to divert QRISs from their primary mission of improving child outcomes. Rating systems that add such elements as the development of a business plan may find that the connection between ratings and important outcomes such as a child's math skills are reduced, if the ratings place too much weight on such elements that presumably make little contribution to children's learning.

Early childhood experts in states and localities designed first-generation QRISs. In building rating rubrics, they relied on their years of experience in ECE classrooms and administration, shared views about best practice in ECE, and on the available research on the factors that promote improved child outcomes. But in many instances, there were no definitive findings about the dimensions of quality that are most important for children's development or the best ways to measure them. Consequently, evidence-based guidance was not always available to help QRIS developers structure their rating rubrics. Furthermore, given the systems' newness, there was no available research or guidance about how to translate dimensions of quality into data that can be collected at reasonable costs, how to assign different ranges of a measure to different quality levels, or how to weight the different elements in creating a simple rating.

This development process led to some similarities across QRISs, but also many differences.⁵ With respect to similarities,

Most QRISs are voluntary, meaning that licensed ECE programs are encouraged (often through incentives) to participate in the rating system but not required to do so.

most QRISs are voluntary, meaning that licensed ECE programs are encouraged (often through incentives) to participate in the rating system but not required to do so. In addition, the majority of QRISs produce ratings on a one-to-five scale, and most link levels of subsidy for serving children of low-income families to a program's rating level (known as tiered reimbursement). And, most commonly, states consider a program rating to be valid for three years, although some rerate as often as annually.

However, beyond these widely adopted features, QRISs vary in substantively important ways: in which quality indicators are included in the rating rubric, how those indicators are combined to produce a summary rating, and in the variety and intensity of quality improvement supports offered to assist programs. For example, in terms of indicators, although nearly every QRIS includes in its program rating some aggregate level of staff education, and nearly all measure aspects of the classroom environment, only a few include a quality standard for whether the program has a written curriculum.

What We Know from Research on QRISs

As they represent a major and relatively new policy, it is important to understand how well QRISs are working and whether they

are achieving their goals. Thanks in part to evaluation requirements specified under federal RTT-ELC grants, there is considerable research under way or recently completed about QRISs.⁶ For example, of the 37 QRISs described in the comprehensive QRIS Compendium that summarizes QRIS features across states,⁷ 33 are reported to have an evaluation under way or completed.⁸ Much of the evaluation work, beginning with RAND's study of Colorado's QRIS,⁹ has focused on assessing whether the rating rubric, which is a central feature of every QRIS, is valid. In addition to validation studies, research on the effectiveness of specific quality improvement supports can help inform QRIS development. Thus, a growing body of research examines the effectiveness of different types of quality improvement supports often found in QRISs. We discuss each of these strands of research in turn.

QRIS Validation Studies

Validation of a QRIS is a multistep process that assesses the degree to which design decisions about program quality standards and measurement strategies produce accurate and meaningful program summary ratings.¹⁰ QRIS validation provides designers, administrators, and stakeholders with crucial data about how well the rating structure and process are functioning. Such studies are very important, given that accurate and meaningful ratings are the basis for

quality improvement support activities, tiered reimbursement, and parental choices. Valid ratings have become even more important, given the substantial public resources that have been flowing into QRISs and the increasing financial stakes associated with rating outcomes (e.g., through tiered reimbursement).

Validation studies typically address two questions: whether programs with a higher rating have higher quality, and whether children in higher-rated programs show better developmental outcomes than children in lower-rated programs.¹¹ To answer the first question, researchers examine whether programs that attain different ratings are ranked in the same way when measured by an independent measure of quality not used in the rating system itself.¹² That is, does a program rated three stars rate lower on the independent measure of quality than a program rated four stars? To answer the second question, researchers examine whether children in higher-rated programs experience greater child development gains than children receiving care in lower-rated programs.

The QRIS Compendium notes that at least one of these approaches has been used in the majority of the 37 QRISs reviewed. Published findings from 12 validation studies covering QRISs in 11 states—California, Colorado, Delaware, Indiana, Maine, Minnesota, Missouri, Oklahoma, Pennsylvania, Virginia,

Valid ratings have become even more important, given the substantial public resources that have been flowing into QRISs and the increasing financial stakes associated with rating outcomes.

and Wisconsin—addressed one or both of these validation questions, and several more such studies are currently under way.¹³

Overall, the findings from the 12 published validation studies we reviewed suggest some progress in the development of valid rating systems, but the evidence is still quite limited and often contradictory, preventing firm conclusions about the validity of QRIS ratings as currently designed. In the ten studies across nine states using the first validation approach described earlier in this perspective, we see that relationships between summary QRIS ratings and one or more independent measures of program quality are positive and significant but modest for most QRISs examined.¹⁴ But differences in quality from one rating level to the next are typically small and may not be linear. Moreover, questions remain about how ratings might vary if rating elements were combined in different ways. Limited research suggests that using different combining rules across the rating elements affects programs' summary ratings.¹⁵

Likewise, of the seven studies that used the second validation approach, examining whether children in higher-rated programs

Overall, the findings from the 12 published validation studies we reviewed suggest some progress in the development of valid rating systems, but the evidence is still quite limited and often contradictory, preventing firm conclusions about the validity of QRIS ratings as currently designed.

show better developmental outcomes than children in lower-rated programs, several have found a modestly positive relationship between QRIS ratings and children's developmental gains in at least one domain during their time in the program. However, these relationships are mostly weak and not found across multiple developmental domains. At the same time, two of the seven studies failed to find any significant relationship between program ratings and children's learning.¹⁶

Several limitations have hampered QRIS validation studies. Studies have been conducted on systems still under development and often include only programs that were among the first to volunteer to participate in the QRIS. These programs may not be representative of the full distribution of programs and likely are more motivated to improve quality. A related issue is that some studies include relatively low numbers of programs that are clustered within just a few rating levels, which limits the ability to detect significant differences across the full range of rating levels. Moreover, parents choose programs for different reasons, so that parental choice of ECE programs may make it difficult to untangle the contribution of the ECE program to child development over and above the contribution that families make. Finally, validation studies that use the first approach to ask whether programs that attain different ratings are ranked in the same way when measured by an independent measure of quality not used in the rating system itself are hampered by the limited number of measures of program quality that are not already used in ratings rubrics.

Research on Quality Improvement Supports

In addition to validation studies, a growing body of research examines the effectiveness of different types of quality improvement

Quality improvement supports can be categorized into four primary types that target individuals: coaching or mentoring, workshops and training, credit-bearing courses at colleges and universities, and formal peer-support activities.

supports often found in QRISs. Ongoing professional learning is viewed as an important activity for ECE staff, many of whom enter their positions without much formal education or training.¹⁷ Quality improvement supports can be categorized into four primary types that target individuals: coaching or mentoring, workshops and training, credit-bearing courses at colleges and universities, and formal peer-support activities.¹⁸ An additional type of support is financial incentives, either at the individual level (e.g., incentivizing staff engagement in quality improvement activities) or at the program level (e.g., tiered subsidy reimbursements to provide additional funds to higher-rated programs).

In general, among these supports, an emerging evidence base suggests coaching efforts are a promising means to improve classroom quality and child outcomes;¹⁹ meanwhile, little to no evidence supports brief workshops and training activities as effective ways to improve classroom or program quality.²⁰ Although coaching shows promise, we lack evidence about what specific features of coaching are most beneficial, such as number of hours or content covered. At the same time, the individualized or classroom-based coaching assistance is relatively resource-intensive compared with workshops that can be offered in large group settings. Additional research is needed to understand whether there are some situations in which workshops might be effective as quality improvement supports,

such as in conjunction with individualized feedback or as part of a structured series of workshops.²¹ Formal peer-support activities may be a promising approach to improve quality at relatively low cost, but these types of supports have not been well studied in the ECE field to date.²² Likewise, the specific benefits of completing credit-bearing courses are unclear because the existing research evidence is mixed.²³

A limitation of this research is that it is often difficult to tease out which specific quality improvement supports are most effective and thus should be prioritized in QRIS design. Much of the existing research lacks randomization into an activity, potentially biasing results because of the self-selection of those who participate (e.g., more motivated staff may choose to participate in a given quality improvement initiative). Systems often bundle two or more support activities together, making it difficult or impossible to disentangle the effects of one versus the other.²⁴ The lack of studies around peer-support activities, which likely reflects both their informality and low numbers of participants, also limits our current knowledge of their effectiveness relative to other supports. Additional research using rigorous designs would contribute to decisionmakers' ability to strategically invest in the most promising quality improvement efforts to support the overall system.

Other Research Gaps

Our review of the QRIS evidence base identifies several gaps in existing research that should be addressed for a more complete understanding of QRIS functioning and effectiveness. First, much of the research on the validity of QRIS ratings and on quality improvement supports has focused on center-based providers; much lower numbers of home-based providers typically are included in such studies. This is, in part, a result of the fact that fewer home-based providers are enrolled in voluntary QRISs in general and fewer still may elect to participate in research studies. We also see that studies include more preschoolers and toddlers than they do infants or school-age children. For the former, this is likely related to fewer numbers of infant classrooms; for the latter, the RTT–ELC validation research required a focus on children prior to school entry.

Second, the validation studies of first-generation QRISs have highlighted important measurement issues with the rating rubrics and are likely to improve the validity of the ratings as a result. But even when rating rubrics are fully vetted and determined to be valid, it will take time to determine whether QRISs, as a policy intervention, are successful at achieving the desired outcome: improved child development outcomes, such as school readiness. The literature on implementing policy reforms makes clear that it takes time to fully implement change, and, even when fully imple-

mented, new systems need time to realize their full effects.²⁵ Once QRISs have demonstrated and implemented valid rating rubrics, evaluation of the effect of QRISs on child outcomes, or other potential areas of impact, is appropriate as part of a larger continuous quality improvement effort.²⁶

Continuing Challenges for QRISs

The past two decades have provided useful information about the design and operation of QRISs and have exposed several challenges with the QRIS approach, with implications for current and future systems. Next, we briefly discuss some of the most significant challenges facing QRISs that are important to consider moving forward.

First, QRISs are designed to improve ECE program quality, with the goal of exposing more children to responsive and developmentally appropriate care and education. To rate program quality, the standard approach—much like ratings of hospitals based on readmission rates or report cards for schools based on children’s standardized test scores—would be to measure the outcomes that ECE programs produce. In other words, for each ECE program, the developmental gains experienced by children in the program would need to be measured, and those results would be used to rate programs according to their ability to promote children’s developmental outcomes. However, measurement of child outcomes is a

The literature on implementing policy reforms makes clear that it takes time to fully implement change, and, even when fully implemented, new systems need time to realize their full effects.

difficult, costly process.²⁷ Young children have to be assessed individually, and sometimes they cannot be coaxed into participating in assessments. From necessity, QRISs turned their attention early on to measuring inputs to quality rather than children's developmental outcomes. Therefore, an inherent challenge with ECE QRISs is the need to identify the appropriate inputs to measure and how much weight to give each quality indicator in constructing the rating rubric.

Second, empirical evidence is limited regarding the dimensions of quality that are most important for children's development and how to translate such dimensions into data that can be collected at reasonable costs. Consequently, evidence-based guidance is not always available to help QRIS developers structure their rating rubrics. In particular, child development experts have concluded that quality in ECE settings is multidimensional and includes a wide range of both structural and process inputs. However, the research regarding some structural indicators—assumed to describe features of the environment that support high-quality care—may be inconclusive or nonexistent. For example, the research is mixed as to whether the education level of the teacher is an important quality feature that predicts child development.²⁸ Rather than the level of education, it may be more important to assess the content and quality of the degree program a teacher attended or the competencies of the teacher as she works with the young children in her care. Likewise, although some measures of process quality that capture teacher-child interactions have shown promise as an input that relates to child development outcomes, other evidence suggests that these measures may not be as predictive as once thought.²⁹ In general, research linking ECE program inputs to child outcomes is hampered by a reliance on observational data that limit the ability

to infer cause-and-effect relationships and can lead to inconclusive findings. As noted earlier, given that parents choose which program to enroll their child in, it is difficult to disentangle the contribution of ECE program inputs from parental inputs in generating the developmental gains that are observed.

Third, structural and, particularly, process measures of quality are often costly to collect, so the focus on inputs in a rating rubric may still require considerable expense to achieve valid and reliable measurements. For example, by their nature, measuring the quality of teacher-child interactions is very time-consuming. Measuring aspects of process quality requires a trained observer to be in the classroom or group setting for relatively long periods of time, which quickly escalates the costs of conducting program ratings. Likewise, verifying staff education and training experiences, observing staff-child ratios, or determining that a curriculum is implemented effectively can involve considerable time on the part of raters. Indeed, the costs of assessing program quality to generate a rating threaten to usurp a large share of QRIS resources.

Given these issues, it is not surprising that researchers investigating optimal rating designs have found it challenging to construct ECE rating rubrics that are strongly predictive of children's development and learning, based on the measures of structural and process quality currently available.³⁰

Furthermore, other aspects of the larger ECE system may not operate as the QRIS logic model assumes, even once a valid rating rubric is put in place. For example, communities may lack a sufficient number of high-quality providers to offer parents meaningful choices, and parents may not be able to afford the high-quality programs that are available. Providers in markets with limited supply may decide, if the QRIS system is voluntary, to forgo participation,

The research base indicates coaching is a promising quality improvement approach, although coaching requires significant investments.

knowing that they will fill their rosters regardless of their quality level. Other providers that are motivated to improve may lack the resources to do so. Providers that improve in quality may not be able to recover their costs, either from private-paying families or via public subsidies.³¹

Recommendations for the Second Generation of QRISs

QRISs have placed a spotlight on the critical issue of ECE program quality. Their development and implementation have engaged diverse stakeholders in the process of quality improvement in ECE settings. They have played and can continue to play an important role in improving the quality of ECE in the United States. At the same time, implementing a complex rating scheme can be costly and may divert resources away from quality improvement efforts and from direct care provision. Furthermore, recent validation studies suggest there is room to improve the rating rubrics that these systems employ. This is a good time to step back and synthesize QRIS evaluation results to date and use them to support a second-generation QRIS design. Although revising a system incurs costs, it is important that ratings have been demonstrated to accurately distinguish lower-, moderate-, and higher-quality providers.

In addition, strategic modifications to existing QRISs may produce savings if they reduce the scope of the required data collection or otherwise streamline the QRIS.

With the goal of moving toward the next generation of QRISs, we offer several recommendations. Those that we direct to state policymakers apply equally well to decisionmakers implementing QRISs at a local level. We believe one of the strongest needs for second-generation QRISs is a focus on direct quality improvement support activities. In addition, there are still many aspects of ratings, including measures of quality and improvement of rating rubrics, that also need research and policy attention.

Recommendation: QRIS Designers Should Direct More Resources to Evidence-Based Quality Improvement Supports.

An increased focus on the “I” in QRIS is critical, particularly for classroom teaching, while researchers accumulate additional evidence needed to construct valid rating rubrics (see next recommendation). States should encourage the use of evidence-based approaches to quality improvement and evaluate the effectiveness of promising new approaches, all with an eye toward identifying the most cost-effective investments. For example, Washington state has studied an ECE coaching program using a randomized controlled trial,³² and other states could undertake similar efforts. As noted earlier, the research base indicates coaching is a promising quality improvement approach, although coaching requires significant investments. Brief stand-alone workshops and training events for ECE staff are most commonly used for professional development and are less costly than individualized coaching yet appear to be least helpful in improving teachers’ classroom practice. Peer-support activities show some promise and could be less costly than

coaching, but this approach needs further evaluation to identify the most-effective models.

Recommendation: The Federal Government or Multistate Consortia Should Direct More Resources Toward the Development of Improved Quality Measures, Which Should Anchor Revised QRIS Rating Structures.

One measurement goal should be the development and identification of better and less-costly measures of classroom or program quality. This includes improved scales that measure multiple dimensions of structural or process quality, as well as more-sensitive indicators of specific quality dimensions. For example, QRISs uniformly assess teacher education level as a proxy for responsive caregiving and overall teaching competence, but the evidence indicates that the relationship is a weak one. Moreover, measuring teacher education level can be costly, as transcripts must be reviewed and equivalencies determined. Instead, measures that focus directly on teacher skills and competencies may be more relevant for children's learning. Such observational measures of classroom interactions and processes are generally quite costly, but technology may play a role in supporting the collection of such evidence (e.g., via videotaping) that is later reviewed and even possibly machine-coded.

Recommendation: States Should Streamline QRIS Rating Rubrics to Focus on Fewer but Most Relevant Aspects of Quality.

Many rating processes are complex, and this complexity is not justified by current research. Each rating item carries a cost associated with measuring it and of altering professional practice to meet it. We suggest that in the absence of definitive evidence, QRISs should focus on fewer rating levels (e.g., two- or three-tier rating systems, instead of five-tier ratings) with fewer but more challenging criteria, such as the adoption and successful implementation of a written curriculum (from a menu of prescreened developmentally appropriate evidence-based curricula) and the use of child assessments to develop individualized learning plans. Until there are improved measures of quality that strongly relate to children's developmental outcomes, it is sensible to keep QRISs focused on the core care and education elements that experts judge to be critical for high-quality ECE programs. Again, there may be upfront costs involved in further modifying a QRIS, but if the result is a more streamlined approach, there could be cost savings down the road that outweigh the costs of a revision.

To identify the most efficient and effective rating structure, researchers could use meta-analytic or pooled-data techniques to synthesize data collected through an array of validation studies recently completed or under way. These data and other second-

There may be upfront costs involved in further modifying a QRIS, but if the result is a more streamlined approach, there could be cost savings down the road that outweigh the costs of a revision.

ary data related to ECE settings could be used to identify quality elements that can be omitted from rating rubrics, where to set cut points on quality measures across rating levels, and more.

Recommendation: States Should Invest in the Full Cost of Implementing and Sustaining the Quality Improvements Specified in the Rating System.

Even if a valid rating system is implemented, public funding is needed to support the one-time costs that providers face to improve quality and the ongoing costs required to deliver high-quality programs, beyond the quality improvement supports referenced in the first recommendation. Most QRISs, for example, offer grants to providers to support the investments in materials, equipment, and facilities needed to improve their quality ratings. QRIS administrators should ensure that those funds are adequate to achieve necessary improvements. Likewise, many QRISs have implemented tiered reimbursement such that reimbursement levels under their

subsidized child care system, as well as state- and locally funded preschool programs, are tied to program rating levels. It is important to ensure that the baseline reimbursement levels are adequate to fund the cost of meeting basic licensing requirements and that the incremental reimbursement tied to each rating level is adequate to cover the cost of the higher quality. If reimbursement levels are not adequate, providers may not be willing to accept families with child care subsidies, thereby limiting access to high-quality ECE programs on the part of lower-income families who could otherwise not afford the full cost of those programs.

The QRIS logic model seeks to achieve a marketplace for ECE where all providers deliver programs in the high-quality range defined by the rating system. To sustain that outcome, providers must be able to recover the costs of providing high-quality programs, and families—through some combination of their own resources and public subsidies—must be able to afford a high-quality program.

NOTES

¹ Jack P. Shonkoff and Deborah A. Phillips, eds., *From Neurons to Neighborhoods: The Science of Early Childhood Development*, Washington, D.C.: The National Academies Press, 2000.

² National Center for Early Childhood Quality Assurance, *Trends in Child Care Center Licensing Regulations and Policies for 2014*, Fairfax, Va.: Department of Health and Human Services, 2015. As of July 1, 2016: https://childcareta.acf.hhs.gov/sites/default/files/public/center_licensing_trends_brief_2014.pdf

³ Gail L. Zellman and Michal Perlman, *Child-Care Quality Rating and Improvement Systems in Five Pioneer States: Implementation Issues and Lessons Learned*, Santa Monica, Calif.: RAND Corporation, MG-795-AECF/SPF/UWA, 2008. As of February 15, 2017: <http://www.rand.org/pubs/monographs/MG795.html>.

⁴ The RTT-ELC grants ranged in size from \$30 million to \$75 million per state for a total of more than \$1 billion awarded between 2011 and 2013. In alphabetical order, the 20 states receiving an RTT-ELC grant and the phase in which they were awarded are as follows: California (phase 1), Colorado (phase 2), Delaware (phase 1), Georgia (phase 3), Illinois (phase 2), Kentucky (phase 3), Maryland (phase 1), Massachusetts (phase 1), Michigan (phase 3), Minnesota (phase 1), New Jersey (phase 3), New Mexico (phase 2), North Carolina (phase 1), Ohio (phase 1), Oregon (phase 2), Pennsylvania (phase 3), Rhode Island (phase 1), Vermont (phase 3), Washington (phase 1), and Wisconsin (phase 2). See U.S. Department of Education, “Race to the Top—Early Learning Challenge,” fact sheet, 2013. As of March 9, 2017: <http://www2.ed.gov/programs/racetothetop-earlylearningchallenge/2013-early-learning-challenge-flyer.pdf>

⁵ BUILD Initiative, QRIS Compendium, homepage, undated. As of December 21, 2016: <http://qriscompendium.org/>

⁶ U.S. Department of Education, “Race to the Top—Early Learning Challenge: Awards,” undated, last updated March 7, 2017. As of March 9, 2017: <http://www2.ed.gov/programs/racetothetop-earlylearningchallenge/awards.html>

⁷ The QRIS Compendium is a web-based database with information about the features of all state and local QRISs gathered by the QRIS Network (see BUILD Initiative, QRIS Compendium, undated).

⁸ We included a single QRIS for California because it operates in many counties and follows a common rating structure, but we did not include the three distinct county-level QRISs operating in Florida.

⁹ Gail L. Zellman, Michal Perlman, Vi-Nhuan Le, and Claude Messan Setodji, *Assessing the Validity of the Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child-Care Quality*, Santa Monica, Calif.: RAND Corporation, MG-650-QEL, 2008. As of February 15, 2017: <http://www.rand.org/pubs/monographs/MG650.html>

¹⁰ Gail L. Zellman and Richard 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, D.C.: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2012.

¹¹ Zellman and Fiene, 2012. A number of other measurement questions typically are addressed in validation studies. For example, do such rating elements as staff education and classroom environment make independent contributions to a program’s rating? They may also address important measurement issues that are often overlooked, such as how to assign measures to rating levels. For instance, must the lead teacher in each classroom have a bachelor’s degree for the program to attain the highest rating, or is an associate’s degree in ECE as good?

¹² Zellman and Fiene, 2012.

¹³ Our synthesis is based on ten studies reviewed as part of RAND’s evaluation of Delaware’s QRIS, as well as findings for Delaware reported in the same study and for California in a separate study. See Lynn A. Karoly, Heather L. Schwartz, Claude Messan Setodji, and Ann C. Haas, *Evaluation of Delaware Stars for Early Success: Final Report*, Santa Monica, Calif.: RAND Corporation, RR-1426-DOEL, 2016. As of March 15, 2017: http://www.rand.org/pubs/research_reports/RR1426.html; and American Institutes for Research and RAND Corporation, *Independent Evaluation of California’s Race to the Top—Early Learning Challenge Quality Rating and Improvement System: Cumulative Technical Report*, San Mateo, Calif., and Santa Monica, Calif.: American Institutes for Research and RAND Corporation, 2016.

¹⁴ Karoly et al., 2016; and American Institutes for Research and RAND Corporation, 2016.

¹⁵ American Institutes for Research and RAND Corporation, 2016.

¹⁶ Zellman et al., 2008; and Katherine Magnuson and Ying-Chun Lin, *Validation of the QRIS YoungStar Rating Scale, Report 2: Wisconsin Early Child Care Study Findings on the Validity of YoungStar Ratings for Children's School Readiness*, Madison, Wis.: University of Wisconsin, Madison, School of Social Work and Institute for Research on Poverty, 2016. As of March 9, 2017: <https://dcf.wisconsin.gov/files/youngstar/pdf/validationreport2.pdf>

¹⁷ Institute of Medicine and National Research Council, *Transforming the Workforce for Children Birth Through Age 8: A Unifying Foundation*, Washington, D.C.: The National Academies Press, 2015.

¹⁸ Peer support activities include formal arrangements such as learning communities, peer support networks, or reciprocal peer coaching to exchange information, and strategies for professional development and program improvement (American Institutes for Research and RAND Corporation, 2016).

¹⁹ Donna M. Bryant, Patricia W. Wesley, Margaret Burchinal, John Sideris, Karen Taylor, Carla Fenson, and Iheoma U. Iruka, *The QUINCE-PFI Study: An Evaluation of a Promising Model for Child Care Provider Training: Final Report Executive Summary*, Chapel Hill, N.C.: Frank Porter Graham Child Development Institute, September 2009; Tabitha Isner, Katherine Tout, Martha Zaslow et al., *Coaching in Early Care and Education Programs and Quality Rating and Improvement Systems (QRIS): Identifying Promising Features*, Washington, D.C.: Child Trends, 2011; and Andrew J. Mashburn, Jason T. Downer, Bridget K. Hamre, Laura M. Justice, and Robert C. Pianta, "Consultation for Teachers and Children's Language and Literacy Development During Pre-Kindergarten," *Applied Developmental Science*, Vol. 14, No. 4, 2010, pp. 179–196.

²⁰ Rachel E. Schachter, "An Analytic Study of the Professional Development Research in Early Childhood Education," *Early Education and Development*, Vol. 26, 2015, pp. 1057–1085; and Martha Zaslow, Kathryn Tout, Tamara Halle, Jessica Vick Whittaker, and Bridget Lavelle, *Toward the Identification of Features of Effective Professional Development for Early Childhood Educators: Literature Review*, Washington, D.C.: Child Trends, 2010.

²¹ Zaslow et al., 2010.

²² American Institutes for Research and RAND Corporation, 2016.

²³ American Institutes for Research and RAND Corporation, 2016.

²⁴ Nikki Aikens and Lauren Akers, *Background Review of Existing Literature on Coaching*, Washington, D.C.: Mathematica Policy Research, 2011; American Institutes for Research, MDRC, MEF Associates, and Child Trends, *Head Start Professional Development: Design Options and Considerations for an Evaluation of Head Start Coaching*, E. C. Howard and K. V. Drummond, eds., Washington, D.C.: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, 2014; and Isner, Tout, Zaslow et al., 2011.

²⁵ Brian Gill, Laura S. Hamilton, J. R. Lockwood, Julie A. Marsh, Ron Zimmer, Deanna Hill, and Shana Pribesh, *Inspiration, Perspiration, and Time: Operations and Achievement in Edison Schools*, Santa Monica, Calif.: RAND Corporation, MG-351-EDU, 2005. As of March 15, 2017: <http://www.rand.org/pubs/monographs/MG351.html>

²⁶ Karoly et al., 2016.

²⁷ Catherine E. Snow and Susan B. Van Hemel, eds., *Early Childhood Assessment: Why, What and How*, Washington, D.C.: National Research Council, 2008.

²⁸ Lynn A. Karoly, *A Golden Opportunity: Advancing California's Early Care and Education Workforce Professional Development System*, Santa Monica, Calif.: RAND Corporation, MG-1188-PF/BECF, 2012. As of February 15, 2017: <http://www.rand.org/pubs/monographs/MG1188.html>

²⁹ Christina Weiland, Kchersti Ulvestad, Jason Sachs, and Hirokazu Yoshikawa, "Associations Between Classroom Quality and Children's Vocabulary and Executive Function Skills in an Urban Public Prekindergarten Program," *Early Childhood Research Quarterly*, Vol. 28, No. 2, 2013, pp. 199–209.

³⁰ Terri J. Sabol, Sandra Soliday Hong, Robert C. Pianta, and Margaret R. Burchinal, "Can Rating Pre-K Programs Predict Children's Learning?" *Science*, Vol. 341, 2013, pp. 845–846.

³¹ David M. Blau, *The Child Care Problem: An Economic Analysis*, New York: Russell Sage Foundation, 2001.

³² Kimberly Boller, Patricia Del Grosso, Randall Blair, Yumiko Jolly, Ken Fortson, Diane Paulsell, Eric Lundquist, Kristin Hallgren, and Martha Kovac, *The Seeds of Success Modified Field Test: Findings from the Impact and Implementation Studies*, Princeton, N.J.: Mathematica Policy Research, Inc., June 28, 2010.

About This Perspective

Quality rating and improvement systems (QRISs) began at the end of the 1990s and have now been almost universally adopted by states and localities as an important tool to boost early child education (ECE) program quality. QRISs are at a critical point in their development and implementation. A wave of QRIS evaluations, most of which are validation studies, are becoming available, largely funded through three rounds of federal Race to the Top–Early Learning Challenge (RTT–ELC) grants covering QRISs in 20 states. The RTT–ELC grants have also been a key source of funds for the development and expansion of ECE QRISs. As the federal grant funds expire, states will have fewer resources available to operate their QRISs without new sources of funding. States will need to be more strategic about the allocation of funds for and within these systems to achieve their goals of expanding access to and improving the quality of ECE programs. In this perspective, we suggest some ways to accomplish this. We assess what the early childhood field has learned about QRISs as they have been widely adopted and matured, and how the field can strategically move this first generation of QRISs into a second generation.

RAND Ventures

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND Ventures is a vehicle for investing in policy solutions. Philanthropic contributions support our ability to take the long view, tackle tough and often-controversial topics, and share our findings in innovative and compelling ways. RAND's research findings and recommendations are based on data and evidence, and therefore do not necessarily reflect the policy preferences or interests of its clients, donors, or supporters.

Funding for this venture was provided by gifts from RAND supporters and income from operations.

© Copyright 2017 RAND Corporation

About the Authors

Jill S. Cannon is a policy researcher at the RAND Corporation and a core faculty member at the Pardee RAND Graduate School. Her work focuses on examining the relationship between early childhood programs and policies and child outcomes.

Gail L. Zellman is a senior research psychologist at the RAND Corporation with years of experience conducting and leading research on child and youth policy. Her research interests include measurement of child care quality, parent school involvement, education innovation, and early socialization.

Lynn A. Karoly is a senior economist at the RAND Corporation and a professor at the Pardee RAND Graduate School. Her research has examined human capital investments, social welfare policy, child and family well-being, and labor markets.

Heather L. Schwartz is a policy researcher and the Associate Director of RAND Education. She researches education and housing policies intended to reduce the negative effects of poverty on children and families.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.html.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. **RAND**® is a registered trademark.

For more information on this publication, visit www.rand.org/t/pe235.



www.rand.org