

HIGHLIGHTS FROM
THE CUMULATIVE
TECHNICAL REPORT

SEPTEMBER 2016



AMERICAN INSTITUTES FOR RESEARCH®

Heather E. Quick
Laura E. Hawkinson
Aleksandra Holod
Jennifer Anthony
Susan Muenchow



Jill S. Cannon
Lynn A. Karoly
Gail L. Zellman
Susannah Faxon-Mills

Independent Evaluation of California's Race to the Top–Early Learning Challenge Quality Rating and Improvement System

In 2012, California was awarded a competitive four-year federal Race to the Top–Early Learning Challenge (RTT–ELC) grant to develop a locally administered, state-supported Quality Rating and Improvement System (QRIS). The California Department of Education has co-led a state implementation team with First 5 California to support county-based Consortia in developing and implementing the state's QRIS rating criteria. In January 2013, 17 Consortia initiated a QRIS that would expand and strengthen preexisting quality improvement



initiatives in 16 counties. The RTT–ELC grant requires an independent evaluation and validation of the QRIS to determine whether ratings are actually associated with higher program quality and whether they can be used to predict children's developmental gains. To this end, the state contracted with American Institutes for Research (AIR) and its partners at RAND Corporation to conduct this study in 2014 and 2015. The study team collected data in a sample of fully rated sites, including independent observations of classroom quality, a survey of providers about their participation in quality improvement activities, and direct assessments of developmental outcomes of 3- and 4-year-old children. However, given the relatively short funding time frame, the mandated evaluation had to begin before the QRISs were fully implemented. Consequently, the key takeaway messages presented below must be considered tentative and mainly should be used to guide refinement of the system and inform the next stage of evaluation.

Key Takeaway Messages

Implementation of the RTT–ELC QRIS was in an early stage at the time the study began, but significant progress has been made over the course of the system’s development, from 2012 to 2015.

- Consortia exceeded their goals for enrolling sites during the RTT–ELC QRIS grant term. They made significant progress in conducting ratings and supporting quality improvement in participating sites.
- California successfully targeted publicly funded programs serving children with high needs for the earliest implementation of the QRIS.
- Few programs had complete QRIS ratings at the start of the study. Given the focus on prioritizing enrollment for publicly funded programs (which must meet minimum quality standards), there was limited variability in program scores among the rated sites; most sites received a score of 3 or 4. These factors limit the applicability of the study findings to the broader set of programs that currently participate in the QRIS.

The dissemination of QRIS ratings has been limited, but analyses of the ratings as well as community input suggest that providing detailed quality element scores may be beneficial.



- As of summer 2015, ratings were used internally for quality improvement and planning purposes but were not yet publicly available to parents in most counties. However, counties had plans to release them, and according to the state implementation team, they did so by the end of 2015.
 - Parents and providers who participated in the study generally agree that the ratings capture the right information about program quality, and parents are eager to have access to the ratings, including information about each measured aspect of quality.
- Analyses of rating elements and their relationship to overall program ratings indicate that QRIS ratings do not represent a single dimension of quality; programs with the same QRIS rating had very different element score patterns. The best information about quality comes from providing element ratings in addition to the overall program rating.

The study provides some evidence of the validity of California’s QRIS ratings, though it is too early in the system’s implementation to draw many conclusions.

- California QRIS ratings are positively related to the quality of classroom interactions in early childhood programs, at least for the limited sample of sites with full ratings. Specifically, higher rated programs were observed to have higher scores on independent measures of the types of teacher–child interactions that are most supportive of children’s developmental outcomes.
- Among programs participating in the QRIS, children had higher average scores on measures of literacy, mathematics, and executive function at the end of the year than at the beginning of the year. This was true for each QRIS rating level. However, as might be expected with the small number of fully rated sites, the limited range in QRIS ratings, and the different populations served by programs at different rating levels, the study found only a small positive relationship between tier ratings and executive function (one of four child outcome measures examined). On measures of early mathematics and literacy skills, analyses did not reveal larger gains

among children attending higher rated programs (Tier 4 or 5), compared with children in Tier 3 programs. Comparisons could not be made with lower rating tiers because of a lack of lower rated programs. These mixed results are consistent with other states' QRIS evaluations.

- Using a slightly different approach to calculating ratings—averaging (instead of summing) scores on the individual quality elements—the study did find slightly stronger relationships with child outcomes. That is, children in sites rated at Tier 4 or Tier 5 using this approach show stronger mathematics and literacy skills at the end of one program year than children in Tier 3 sites.
- These results cannot be used to draw firm conclusions about the validity of the system given its early stage of implementation. Further evaluation once the system is more mature and programs representing a wider ratings distribution are enrolled will be necessary to draw conclusions about the relationship between attending a higher rated program and children's developmental outcomes.
- Study results also cannot be used to assess the causal relationship between program quality and child outcomes, given the observational nature of the study design. To do so would require an experimental approach in which children are randomly assigned to programs with different levels of quality. This design was not possible for this study, whose primary purpose was to assess the validity of the rating system and the status of the improvement activities.

Study analyses reveal high levels of participation in quality improvement activities by program staff and point to coaching as a promising approach to improving quality.

- Large numbers of program staff in rated sites reported participating in a range of quality improvement and supportive activities, including coaching and mentoring, workshops and training, peer supports, and credit-bearing courses.
- Staff reported that coaching was the most helpful strategy for their professional learning, although coaching is relatively expensive compared with other types of quality improvement supports.
- In addition, the study found evidence that more intensive or sustained coaching is positively linked to program quality and children's developmental outcomes.



Limitations of the study should be considered in interpreting the results.

First, the sample of programs participating in the study was smaller than expected

and had limited variability in ratings and program characteristics, so results of these analyses might differ from a representative sample of all types of programs in the state. In addition, the exploratory design of the study does not provide information about the causes underlying the observed changes in program quality and children's developmental outcomes; additional unobserved factors could be driving these changes. Nonetheless, findings from the pilot phase of California's QRIS can be used to inform decisions about refinements to the system for wider implementation.

Policy Options for Consideration

Our analyses suggest some directions that may be worth consideration by the state.

1. **Consider presenting detailed rating information to parents that includes program ratings as well as element scores.** The state and Consortia also should consider conducting a pilot phase for releasing rating data to the public to identify the best communication strategies.
2. **Consider alternative rating strategies, such as taking an average score across elements to strengthen validity.** Other approaches, such as modifying cut points on the element scores, also are worth exploring.
3. **Consider a coaching model that offers sustained support over the course of the year.** Sustained work with a coach over more hours during the course of the year appears to promote positive outcomes.
4. **Consider exploring the types of peer supports that are available to staff to learn more about effective approaches to this type of quality improvement activity.** Observed associations between classroom quality and participation in peer support networks were present but weak; it would be valuable to examine what these interactions entail and under what conditions they appear to be more or less effective.
5. **Consider ways to encourage or require a broader range of providers to participate.** The lack of variation in rated quality limits the power of evaluation studies, reduces the potential impact of the QRIS in helping families choose programs, forgoes opportunities to assess the quality of the large group of private programs receiving some public voucher funds, and makes it difficult for the public or policymakers to determine how best to direct limited resources for quality improvement. Pilot studies that require program participation, which occurs in a number of states, would produce useful data about a more inclusive system.
6. **Consider another validation phase once the system is refined and expanded.** In such a study, it would be essential for Consortia to collect, maintain, and share with the state additional classroom- and site-level data. Such data would enable additional analyses and suggest evidence-based refinements not possible without these more detailed data.

Funding for the study was provided by the California Department of Education.

We also acknowledge the contributions of Deborah Parrish, Alejandra Martin, Emily Weinberg, and Dong Hoon Lee from AIR; and Ashley Muchow and Tiffany Tsai from RAND.

For more information, contact Heather Quick at hquick@air.org or view the full Executive Summary and Cumulative Technical Report at <http://www.cde.ca.gov/sp/cd/rt/rttelc.asp>.

About AIR

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

AIR's early childhood development research focuses on evaluating programs and policies, improving professional development, examining accountability and assessment systems, investigating program quality and classroom practices, and translating research to practice to aid young children and their families.



1000 Thomas Jefferson Street NW, Washington, DC 20007-3835
877.322.8700 | 202.223.6690

7299_09/16