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

Cumulative Technical Report

Submitted to:

California Department of Education
Early Education and Support Division

Submitted by:

American Institutes for Research
RAND Corporation

AIR®
AMERICAN INSTITUTES FOR RESEARCH®

RAND CORPORATION

August 2016

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

Cumulative Technical Report

August 2016

Project Leadership:

Heather E. Quick, *Project Manager*
Laura E. Hawkinson, *Analysis Lead*
Aleksandra Holod, *Analysis Lead*
Susan Muenchow, *Senior Advisor*
Deborah Parrish, *Senior Advisor*
Jill S. Cannon, *RAND Project Manager*
Susannah Faxon-Mills, *RAND Deputy Project Manager*
Lynn A. Karoly, RAND *Senior Advisor*
Gail L. Zellman, RAND *Senior Advisor*

Report Authors:

AIR team: Heather E. Quick, Laura E. Hawkinson, Aleksandra Holod, Jennifer Anthony, Susan Muenchow, Deborah Parrish, Alejandra Martin, Emily Weinberg, and Dong Hoon Lee

RAND team: Jill S. Cannon, Lynn A. Karoly, Gail L. Zellman, Susannah Faxon-Mills, Ashley Muchow, and Tiffany Tsai

Allen, Shea & Associates team: Mechele Small Haggard

AMERICAN INSTITUTES FOR RESEARCH®

2800 Campus Drive, Suite 200
San Mateo, CA 94403
650.843.8100 | TTY 877.334.3499
www.air.org

RAND CORPORATION

1776 Main Street
Santa Monica, CA 90401-3208
310.393.0411
www.rand.org

Copyright © 2016 American Institutes for Research. All rights reserved.

Acknowledgments

The American Institutes for Research and RAND Corporation study team wish to extend our sincere appreciation to the representatives from the 17 RTT-ELC Regional Leadership Consortia that contributed to this research. In particular, we thank the county administrators who participated in interviews, provided data, and encouraged their early learning and development sites to participate in the study. We also would like to thank the site administrators, providers, and teachers who opened their doors to the study team to share their experience and allow us to observe their practice. We are also grateful for the many study families for their participation in focus groups as well as for allowing their children to participate in the study

We also wish to acknowledge the invaluable assistance of data analysts and research support staff who contributed to this research, including Raquel González, Natalie Tucker-Bradway, Debbie Davidson-Gibbs, Kiana Abram, Kaitlin Fronberg, Carmen Martínez, Christine McGuigan, Nicol Christie, Michele Cadigan, Suzette Chavez, Emily Anderson, Benjamin West, Nicholas Mills, John Meakin, Martha Ramirez, Erik Loewen, Shannon Keuter, Shaheen Khan, John Mezzanotte, Megan Brown, and Robin Gurley. We also thank Anja Kurki, Ann-Marie Faria, Bokhee Yoon, James Elicker, and Cathy Stasz for their technical and content reviews and feedback.

Thanks also go to our partners at Survey Research Management – Linda Kuhn, Tony Lavender, Betsy Quicksall, Aimee Elsey, Lyn Bopp, Ashley Bronzan, Daniel Mackin, and all of the field staff who collected data; and at Allen, Shea & Associates – William Allen and Mony Flores-Bauer for their thoughtful contributions to research support activities.

Last, we wish to acknowledge the guidance and input provided by Cecelia Fisher-Dahms, Channa Hewawickrama, Gretchen Williams, Sarah Neville-Morgan, David Dodds, Debra Silverman, Debra McMannis, Maria Balakshin, Serene Yee and other staff of the California Department of Education, Early Education and Support Division, and First 5 California.

Contents

Acknowledgments i

Chapter 1. Introduction 1
- History and Purpose of QRISs Nationally and in California 1
- QRIS Evaluation and Validation Studies 4
- The Independent Evaluation of California's RTT-ELC QRIS 6

Chapter 2. Implementation of the RTT-ELC QRIS 14
- Status of the Implementation of the RTT-ELC QRIS 16
- The Hybrid Rating Matrix 22
- Workforce, Professional Development, Training, and Technical Assistance 27
- Enhancement and Alignment of Early Care and Education Systems and Initiatives 29
- Next Steps 31
- Summary 43

Chapter 3. Validity of the QRIS Ratings 45
- How Well Does the QRIS Perform as a Measure of Quality? 47
- How Well Does the QRIS Differentiate Programs Based on Observed Quality? 53
- How Well Does the QRIS Predict Children's Outcomes? 56
- How Do Alternative Rating Approaches Affect the Distribution and Validity of Ratings? 68
- How Well Do Alternative Rating Approaches Predict Children's Outcomes? 71
- Summary 75

Chapter 4. Perspectives on Quality Elements and Ratings 79
- Perspectives on Elements of the Hybrid Rating Matrix 80
- Other Elements That Influence Perceptions of Quality and Parent Choices 87
- Publicizing Ratings 90
- Summary 92

Chapter 5. Staff-Level Quality Improvement Supports 93
- Staff Sample Demographics 94
- Coaching or Mentoring Supports 96
- Noncredit Workshops or Training 101
- Peer Support 105
- Credit-bearing College or University Courses 108
- Financial Incentives 111
- Perceptions of Quality Improvement Activities 115
- Summary 118

Chapter 6. Program-Level Quality Improvement Supports 121
- Director and Site Sample Characteristics 123
- RTT-ELC QRIS Participation 124
- Perceptions of Tier Advancement 126
- Site-Level Quality Improvement Support Efforts 130

Director Quality Improvement Supports 135
Summary 138
Chapter 7. Quality Improvement Activities and Changes in Teacher Practice and Children's Outcomes 140
Relationship Between QI Activities and Classroom Quality 143
Relationship Between QI Activities and Children's Outcomes 151
Summary 158
Chapter 8. Cost of QRIS Consortia Quality Improvement Supports 161
Per Unit Cost of QI Supports 163
Summary 166
Chapter 9. Summary and Conclusions 168
Summary of Findings 168
Study Conclusions and Limitations 183
Policy Options for Consideration 186
References 191

Chapter 1. Introduction

California's Race to the Top–Early Learning Challenge (RTT-ELC) grant provided funding for the development of a locally driven Quality Rating and Improvement System (QRIS) or set of systems as well as an independent evaluation to validate the rating approach and assess outcomes associated with participation in the system. In January 2014, the California Department of Education (CDE) contracted with American Institutes for Research (AIR) and its partners at the RAND Corporation; Survey Research Management; and Allen, Shea & Associates to conduct the evaluation. The first year's validation results were presented in the half-term report (http://www.cde.ca.gov/sp/cd/rt/documents/airhalftermreport.pdf).

This final comprehensive report highlights key findings from the half-term report (see chapter 2) and presents additional results related to the implementation of the system, quality improvement (QI) supports provided through the system, program quality and children's developmental outcomes, and perceptions of quality and the rating system.

In this introductory chapter, we present a brief summary of the history and purpose of California's QRIS as well as a review of what other QRIS evaluation studies have found. We provide an overview of the goals and approach used in the evaluation of California's RTT-ELC QRIS, including the study questions and methods that drove the study. This chapter concludes with an overview of the report, its structure, and content.

California's RTT-ELC QRIS

In 2011, California successfully submitted a Race to the Top–Early Learning Challenge (RTT-ELC) grant application to the U.S. Department of Education that would move the state toward a locally driven Quality Rating and Improvement System (QRIS) or set of systems. The state proposed building a network of 17 Early Learning Challenge Regional Leadership Consortia that had already established—or were in the process of developing—QRIS initiatives in 16 counties. These Consortia, comprised of local First 5 commissions, county offices of education, and other key stakeholders, represent counties that together have more than 1.8 million children ages birth to five. This locally based approach sets some common goals for workforce development, program assessment, and child assessment for school readiness, but allows for some flexibility in quality benchmarks. The counties participating in the RTT-ELC Regional Leadership Consortia have voluntarily adopted a Hybrid Rating Matrix that allows considerable local autonomy in some tier requirements, the rating protocol, and supports and incentives for quality improvement.

History and Purpose of QRISs Nationally and in California

Research findings highlight the importance of the period from birth to school entry for child development and focus attention on the quality of care and early learning experiences that young children receive (Center on the Developing Child 2007; National Research Council 2001; Shonkoff and Phillips 2000; Vandell and Wolfe 2000). Numerous studies have demonstrated that higher quality care, defined in various ways, is related to positive developmental outcomes for children, including improved language development, cognitive functioning, social competence, and emotional adjustment (e.g., Burchinal and others 1996; Clarke-Stewart and others 2002; Howes 1988; Mashburn 2008; National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network [ECCRN] 2000; Peisner-Feinberg and others

2001; Weiland and others 2013), although the benefits tend to be largest for children from disadvantaged backgrounds (e.g., Gormley and Gayer 2005; Gormley and others 2005; Karoly 2009; Pianta and others 2009). More recent studies that examine the effects of quality thresholds (whether a particular quality level must be achieved to demonstrate effects on children; summarized in Zaslow and others 2010), which find that associations between quality and outcomes are stronger when observed quality is in the higher range, underscore the importance of high-quality care in improving child outcomes.

Research also suggests that, when faced with choices in early care for their children, parents are not always accurate in rating the quality of care provided (e.g., Helburn, Morris, and Modigliani 2002). Parents tend to rate child care providers very positively (e.g., Barraclough and Smith 1996; Cryer and Burchinal 1997; Helburn 1995; Wolfe and Scrivner 2004), and their ratings do not correlate with trained observer quality ratings (e.g., Barraclough and Smith 1996; Cryer and Burchinal 1997; Cryer, Tietze, and Wessels 2002). Many parents (erroneously) believe that licensing includes scrutiny of program quality and that licensure indicates that a program is of high quality (National Association of Child Care Resource and Referral Agencies 2011).

These findings highlight the need for systematic, reliable, and valid information about the quality of the care and early learning environments for young children—such as that provided through a QRIS—to be publicly available. Thus, QRISs aim to (1) provide quality information to parents to inform their choice of early learning and development programs for their children, and (2) expand meaningful parental choice by supporting program QI.

All but one state currently implement or plan to implement some form of QRIS (QRIS National Learning Network 2015). QRISs have recently garnered national attention through the U.S. Department of Education's (ED's) RTT-ELC grant program. In the RTT-ELC request for applications, ED encouraged each state to design and implement a tiered QRIS that was standards based and provided "meaningful" ratings for the quality of each program. ED also encouraged broad participation in the QRIS across program types, with a priority toward including all licensed or state-regulated early learning and development programs. In addition, ED emphasized a focus on continuous program improvement and a dissemination plan for ratings that would allow families to make informed decisions about which programs could best serve the needs of their children. Also required as part of RTT-ELC funding was a rigorous evaluation and validation of the QRIS (U.S Department of Education, 2011, p. 8).

In California, the movement to create a QRIS predates the federal focus on QRIS development. Beginning in 2004, First 5 California funded Power of Preschool initiatives featuring many of the typical elements of a QRIS: quality standards, provider support, program quality assessments, ratings to determine the level of payment (or subsidy) tied to quality ratings, and financial incentives for QI. A number of counties established their own initiatives designed to use publicly disseminated ratings as the major impetus for QI.

In 2008, Senate Bill 1629 established a California Early Learning Quality Improvement System (CAEL QIS) Advisory Committee to design a QRIS for California. The committee produced a report in December 2010 that detailed a design for a QRIS with a block system (where all elements in one tier must be achieved before advancing to the next tier) that included five quality elements for the rating structure. The CAEL QIS Advisory Committee proposed piloting the system over three years before implementing it on a statewide basis and advised that the system should be phased in over five years or more, after completion of the pilot. In 2011, before the piloting of the proposed system had begun, the State of California—citing serious budget concerns as well as the challenge of implementing a one-size-fits-all program in such a large and diverse state—successfully submitted an RTT-ELC application that moved toward a more locally driven QRIS approach. The state proposed building a network of 17 ELC Regional Leadership Consortia across 16 counties that already had established, or were in the process of developing, QRIS initiatives. Key participants in the Consortia include local First 5 commissions and county offices of education as well as other stakeholders.

California QRIS Key Terms

Consortia: County-based agencies administering the QRIS locally

Tiers: California QRIS rating levels, ranging from 1 (lowest) to 5 (highest)

Elements: Aspects of quality measured in California's QRIS. Programs receive scores from 1 to 5 on as many as seven elements (the number of rated elements depends on the program type). The element scores are used to determine the program's tier.

Hybrid Rating Matrix: The California QRIS document that outlines criteria for each element score, as well as criteria for each tier. Consortia may make local adaptations to the criteria for Tier 2 and Tier 5.

Continuous Quality Improvement Pathways: The California QRIS document that outlines additional aspects of quality that are not measured for the QRIS but are prioritized as part of the state's Quality Continuum Framework.

In 2013, a new QRIS was adopted by 17 Consortia, which include a mix of small and large counties representing diverse areas of the state, as well as some counties with no previous QRIS experience and other counties that had operated local QRISs for as long as a decade. The participating Consortia worked with the CDE to develop the Hybrid Rating Matrix, which specifies the criteria for five rating levels. The Consortia agreed to adopt the rating criteria in the Hybrid Rating Matrix, with the option to make some local adaptations to Tiers 2 and 5 while maintaining three common tiers (Tiers 1, 3, and 4). The California QRIS is referred to as a hybrid rating approach because ratings are determined using a combination of points earned by meeting standards in different quality elements and "blocks" that require programs to meet minimum criteria across elements for a given rating level. The Hybrid Rating Matrix has block requirements for Tier 1 and offers point ranges for Tiers 2, 3, 4, and 5. However, the Consortia have the local option to treat Tiers 2 and 5 as blocks. Other local adaptations to Tiers 2 and 5 include adding supplemental criteria to reach the tier in addition to the blocks or point ranges specified in the Hybrid Rating Matrix.

The QRIS ratings that result from the Hybrid Rating Matrix are intended for multiple purposes. They are expected to be reliable and meaningful and inform parents about program quality, to differentiate programs according to the quality of program structures and adult-child interactions, to inform program quality improvement efforts, and to identify programs that best support child learning and developmental outcomes.

Accompanying the Hybrid Rating Matrix as part of a Quality Continuum Framework are the Continuous Quality Improvement Pathways. The Pathways Core Tools and Resources include the California *Foundations* and *Frameworks*, *Preschool English Learners Guide*, the Desired Results Developmental Profile assessment, Ages and Stages Questionnaire, Center on the Social and Emotional Foundations for Early Learning (CSEFEL), Strengthening Families Protective Factors Framework, and other resources listed in the federal application that the Consortia are required to include in their quality improvement plan. The Consortia are to gather data regarding how these tools and resources are used. Although some of the resources also are listed in the Hybrid Rating Matrix, others are not included in the ratings.

QRIS Evaluation and Validation Studies

The investment of considerable federal and state funds to improve the quality of early learning and development programs using QRIS initiatives has increased the need for informative and rigorous evaluations of QRISs across states. A major component of QRIS evaluations are validation studies that examine properties of program ratings. As a tool, QRISs have tremendous potential to transform the early childhood landscape; however, the utility of QRISs is only as good as the ratings on which they are based. Validation studies determine whether these ratings are accurate measures of quality and, more specifically, whether the QRIS ratings serve as a valid measure for their intended purposes. Validation studies of existing QRISs are needed to demonstrate that ratings within the systems are meaningful and accurate and that they successfully differentiate low-quality programs from high-quality programs. When conducted with rigor, validation studies of QRISs assess whether the ratings developed in the system can be accurate indicators of program quality and whether they predict learning and developmental outcomes for children. In addition to the validation of the rating itself, evaluations of QRISs also are needed to demonstrate that these systems, compared with a counterfactual with no QRIS in place, are effective in raising the quality of early learning programs and improving child outcomes.

The goals of QRIS validation research are different depending on the stage of QRIS development and implementation. Validation research in the early stages of QRIS implementation can be used to inform decisions about revisions to the QRIS rating approach and can lead to different implementation strategies or additional training and supports to ensure successful QRIS implementation as the system expands. This early validation research also can inform later efforts to evaluate the system after the QRIS has been finalized and broadly implemented. Validation and evaluation at later stages, when the system is fully implemented, can provide more definitive information about the properties of the ratings and the effectiveness of the system.

In a literature review for the *Local Quality Improvement Efforts and Outcomes Descriptive Study* (AIR and RAND 2013) and updated for the half-term report, the AIR/RAND study team found that although QRISs are being designed and implemented in most states, evaluation evidence of

QRISs comes from just 12 states or substate areas.[2] Our review of QRIS evaluation studies produced the following key points regarding validation and impact findings (Barnard and others 2006; Boller and others 2010; Bryant and others 2001; Elicker and others 2011; Lahti and others 2013; Malone and others 2011; Norris and Dunn 2004; Norris, Dunn, and Eckert 2003; Sabol and Pianta 2012, 2014; Shen, Tackett, and Ma 2009; Sirinides 2010; Thornburg and others 2009; Tout and others 2010, 2011; Zellman and others 2008):

- The 14 evaluations (across 12 states or substate areas) we identified almost exclusively consist of validation studies that address one or more questions about the effectiveness of the QRIS design in differentiating programs based on quality. Only one study provides any evidence of the causal impact of a QRIS and only for a narrow question (namely, did the addition of coaching, QI grants, and funds for professional development have an effect on staff professional development, observed care quality, and program QRIS ratings?).
- Eleven of the 14 studies examined the relationship between QRIS ratings and a measure of program quality (Barnard and others 2006; Bryant and others 2001; Elicker and others 2011; Lahti and others 2013; Malone and others 2011; Norris and Dunn 2004; Norris, Dunn, and Eckert 2003; Sirinides 2010; Tout and others 2010, 2011; Zellman and others 2008). Ten of the 11 studies used the Environment Rating Scales (ERS) as an outcome measure. All but one found that the system ratings were correlated positively with observed quality, although the correlation was not always statistically significant. Moreover, the ERS was generally not an independent measure of quality, as it was used to determine the ratings that were being validated.
- Six studies aimed to determine whether program ratings or other program quality measures improve over time (Elicker and others 2011; Norris, Dunn, and Eckert 2003; Shen, Tackett, and Ma 2009; Sirinides 2010; Tout and others 2011; Zellman and others 2008). These studies provide consistent evidence, given the way quality is defined, measured, and incentivized in the QRIS, that programs can raise their rating and improve their quality over time.
- Seven studies examined the relationship between QRIS ratings and child developmental outcomes (Elicker and others 2011; Sabol and Pianta 2012, 2014; Shen, Tackett, and Ma 2009; Sirinides 2010; Thornburg and others 2009; Tout and others 2010, 2011; Zellman and others 2008). The findings from these studies are mixed, at best, indicating that there is little evidence to date to suggest that QRIS ratings, as currently configured, are predictive of child gains for key developmental domains.
- Two studies provide validation evidence about parents' knowledge and understanding of the QRIS ratings (Elicker and others 2011; Tout and others 2010). These studies conclude that parents in rated programs know more about the rating system than the general public does and that knowledge of the system tends to increase over time. Even so, the extent of

[2] With the requirement for evaluation as part of the RTT-ELC grants, additional QRIS validation studies have been initiated and have produced or will be producing additional findings beyond those summarized in our latest literature review.

parental awareness of the examined QRISs did not exceed 20 percent for the general public and 40 percent for those using rated providers.

- Although QRIS designers may ultimately be interested in measuring the impact of implementing key elements of an individual QRIS, or QRISs as a whole, on a range of system outcomes—provider mix, parental choice, teacher professional development, program quality, or child outcomes—making such causal inferences requires experimental or quasi-experimental designs that have rarely been implemented to date. The one available experimental study (Boller and others 2010) of enhancements to the QI activities in the pilot of the Washington State QRIS demonstrates the potential for using scientifically rigorous methods to extend our understanding of the causal impacts of QRIS implementation.[3]

The complete literature review can be found in appendix A of the half-term report (http://www.cde.ca.gov/sp/cd/rt/documents/airhalftermreport.pdf).

The Independent Evaluation of California's RTT-ELC QRIS

The evaluation of California's RTT-ELC builds on this prior research on QRISs and focuses on five main aspects of the system: (1) overall implementation, (2) the validity of the ratings, (3) the QI supports that are provided to early learning staff through the system, (4) quality outcomes and children's developmental outcomes for sites participating in the system, and (5) parents' and providers' (teaching staff, center directors, and family child care home [FCCH] providers) perceptions of quality and the rating system. These areas of focus provide the organizational structure for this report, as depicted in exhibit 1.1 and described in more detail below. In addition, each of the areas of focus aligns with one or more research questions guiding the study analyses, as shown in exhibit 1.2. In addition, the analyses line up with the intended purposes of the QRIS: to serve as reliable and meaningful ratings to inform parents about program quality, to differentiate programs according to the quality of program structures and adult-child interactions, to inform program quality improvement efforts, and to identify programs that best support child learning and developmental outcomes. Additional detail on sampling, data collection, and analysis can be found in appendix 1A.

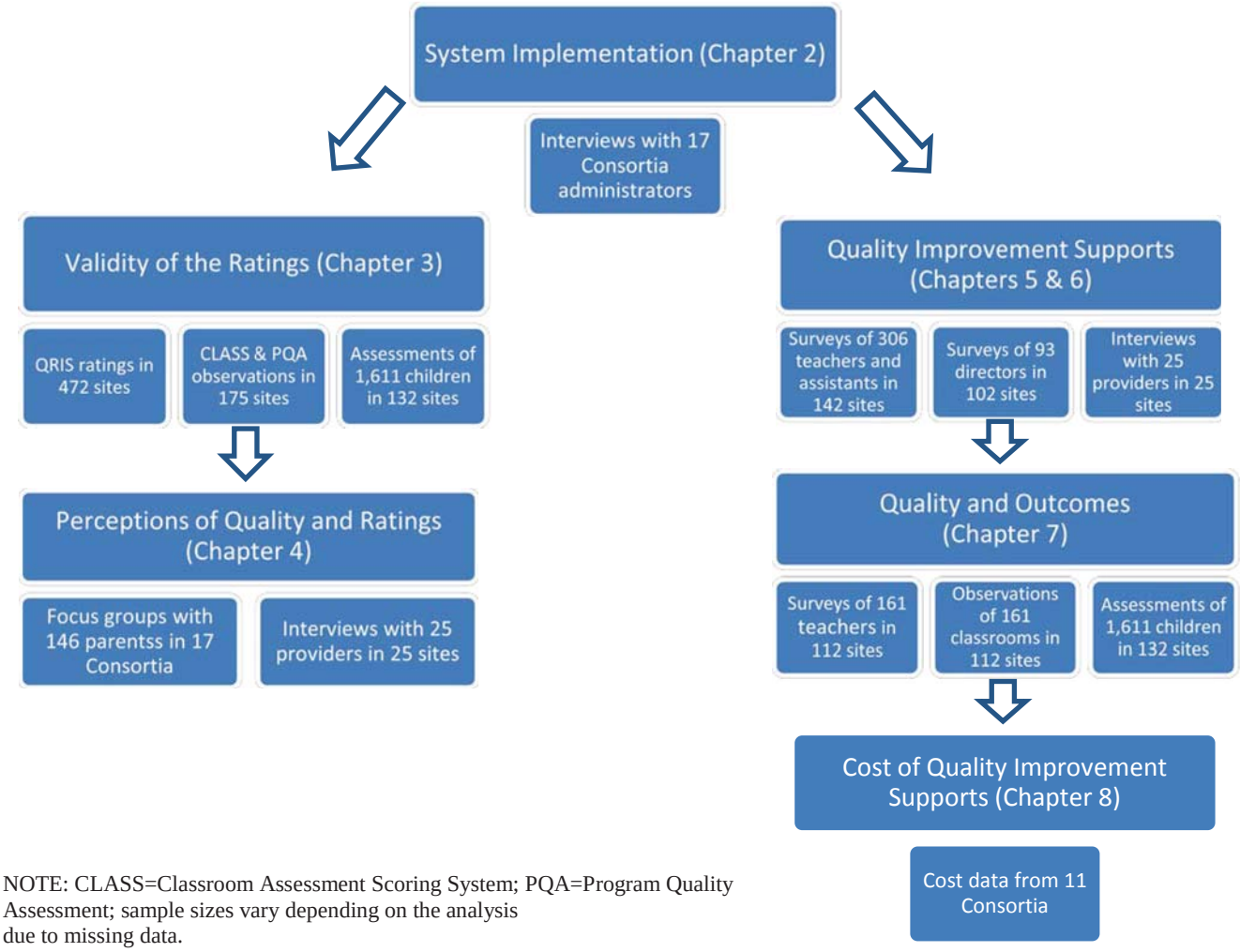
System Implementation

It is important to consider the stage and degree of implementation of a QRIS in its evaluation. In this report, the system implementation study draws on information gathered from interviews that the AIR/RAND team conducted with the administrators of each of the 17 QRIS programs in the Regional Leadership Consortia in spring/summer 2015. These interviews were designed to learn more about the work the Consortia have done on their QI systems over the course of the 2014–15

[3] The experimental evaluation of the Seeds for Success pilot QRIS found that six months after random assignment, compared with the control programs, programs that received the treatment (consisting of coaching, quality improvement grants, and additional funds for teacher professional development) had significantly higher observed quality but no change in program ratings under the Seeds for Success block design (Boller et al., 2010). There also were favorable effects on teacher participation in education and training, on course credits received, and on turnover.

program year. Depending on the structure of the Consortia, AIR and RAND interviewed staff of local First 5 offices, county offices of education, and other key partners. Using qualitative data analysis techniques, the study team analyzed the interview transcripts to gain an understanding of the work of each Consortium and to identify differences and common themes across Consortia. These data were supplemented with interviews conducted with 25 providers who were asked questions about their experiences with the QRIS.

Exhibit 1.1. Structure of the RTT-ELC QRIS Evaluation and Report



NOTE: CLASS=Classroom Assessment Scoring System; PQA=Program Quality Assessment; sample sizes vary depending on the analysis due to missing data.

Validity of the Ratings

As a follow-up to the validation study conducted as the first component of the evaluation in the 2013–14 program year, we report on a broader set of analyses examining the validity of the ratings in this report. We first summarize key findings presented in the half-term report that investigate the extent to which the ratings assigned by Consortia differentiate programs based on observed measures of quality. These analyses draw on data gathered in 2014, including the 2013

ratings data (Common Data Elements) for all 472 programs with full and complete ratings, which enable us to examine the distribution of ratings across centers and FCCHs among all fully rated programs. These data, submitted to the state using the QRIS reporting requirements, include information on program type, enrollment, funding sources, languages spoken in the program, element scores, the sum of the element scores, the QRIS rating, and the program's average Classroom Assessment Scoring System (CLASS) scores used to calculate the CLASS element scores. Data were available for 1,272 programs, although only 472 had full ratings; the remaining 800 did not have full ratings, reflecting the early stage of implementation.

Next, classroom observations were conducted through spring 2014 using the CLASS and Program Quality Assessment (PQA). The study team selected two independent observation instruments in order to compare QRIS ratings to a measure of program quality that is widely used and closely connected to the QRIS system (the CLASS instrument, which is factored in to one of the seven QRIS element scores), and also compare QRIS ratings to another validated measure that is not part of the rating calculation but measures aligned program quality constructs. At the request of several of the Consortia and the CDE, we accepted some extant CLASS data from the Consortia in lieu of conducting direct observations of classrooms if the data had been collected within nine months of the study's data collection period. Classroom observation data were obtained for 175 sites. By comparing CLASS and PQA scores for programs at different rating tiers, we evaluate the extent to which the ratings are successful at discriminating (or distinguishing) among programs that vary in quality on these classroom quality measures.

A second component of validating ratings involves examining the degree to which the ratings differentiate programs based on children's developmental outcomes. To address this question, we conducted direct one-on-one assessments of 1,611 three- and four-year-old children from a sample of 132 fully rated programs in fall and spring of the 2014–15 program year. We used a range of developmental measures, including the Woodcock-Johnson Letter-Word Identification subtest (Woodcock, McGrew, and Mather 2001) and the Story and Print Concepts assessment (Zill and Resnik 2000) to assess preliteracy skills, the Woodcock-Johnson Applied Problems subtest (Woodcock, McGrew, and Mather 2001) to assess mathematics skills, and the Peg Tapping task (Diamond and Taylor 1996) to assess executive function. We compared outcomes for programs rated at different tiers, using the 2014 Common Data Elements to supplement the 2013 ratings data.

Using a combination of the 2013 and 2014 Common Data Elements, we recalculated the ratings using several alternative calculation methods, such as blocking at Tier 2 as compared with the standard RTT-ELC method of applying a block at Tier 1 and assigning points for higher tiers. This was done in order to identify ways to improve the validity of the ratings. Using the new ratings, we examined the quality and child development gains for programs at different tier levels to determine if the alternative approaches differentiate programs better.

Quality Improvement Supports

A primary goal of the QRIS is to improve the quality of early learning and care programs. QRISs attempt to do this through the provision of QI supports, such as training, coaching, and financial incentives. To characterize the supports provided to staff, we administered a survey to 306 staff

over the course of the 2014–15 program year, asking about the range of QI experiences and supports that were provided for center teachers and assistants as well as family child care providers. We also surveyed 93 program directors in spring 2015 to identify program-level supports provided to improve program quality. These data were analyzed to present a picture of the availability and utilization of QI supports across the 11 focal Consortia. These analyses were supplemented with qualitative information from the provider interviews described above.

Quality and Outcomes

A central focus of the evaluation was to increase understanding of the relationships between participation in QI activities and changes in program quality and developmental gains of children attending participating programs. To address this question, we used data on QI participation from the staff survey described above. In addition, to examine quality outcomes, we conducted additional CLASS observations in spring 2015 in 112 sites. We combined these CLASS scores with the CLASS scores obtained in spring 2014 to provide assessments of quality at two points in time. We then used QI participation variables in regression analyses to predict 2015 CLASS scores, controlling for 2014 CLASS scores, and to explore the relationships between QI and quality outcomes.

To examine how children's developmental outcomes differ for teachers who have participated in various types of QI activities, we used the staff survey data, as well as children's fall and spring assessment data, for a subsample of 132 sites, which includes 1,611 children with teacher survey responses. Then, using QI variables and multilevel modeling techniques, we predicted children's spring assessment scores, controlling for fall scores.

Cost of Quality Improvement Activities

The implementation of a QRIS—particularly the QI components of the system such as coaching, mentoring, credit-bearing courses, and so on—can represent a significant investment of resources. To our knowledge, there has been little attention paid to date in QRIS validation studies to the costs associated with QI activities. Such cost data may be of interest in their own right as a way of understanding the value of the resources required to support QI activities of various types. In addition, when combined with estimates of the impacts of QI supports on program QI or children's developmental gains, the cost information can be used to compare the relative cost-effectiveness of each type of QI support.

To measure the cost of QI activities, we gathered data on expenditures and in-kind resources for the 11 focal Consortia specific to the main types of QI supports: coaching/mentoring, credit-bearing courses, noncredit-bearing courses, peer support activities, and financial incentives. Each local Consortium also provided information on the outputs associated with each type of QI support (for example, the number of program staff receiving coaching, the total number of coaching hours provided, and so on). This information was used to calculate estimates of the average cost per unit of QI activity across the Consortia with valid data.

Perceptions of Quality and Ratings

Finally, to better understand the potential for the QRIS to influence parents' understanding and use of information about quality care environments as well as to examine parents' and providers' perceptions of the ratings and the rating system, we analyzed content from the provider interviews described above and from focus group discussions with parents. We conducted 17 focus groups—one in each Consortium—which included a total of 146 parents whose children attended a range of center-based and family child care programs. Parents were asked about their priorities in choosing an early learning and care program and also provided feedback on the quality elements included in the Hybrid Rating Matrix. These data were analyzed using qualitative techniques to identify common themes regarding perceptions of quality and the rating system.

Exhibit 1.2. Evaluation Questions Addressed by Each Study Component

Research Question	System Implemen-tation	Perceptions of Quality and the Ratings	Validity of the Ratings	Quality Improve-ment Supports	Quality and Outcomes
System Implementation					
1. What is the status of implementation of the RTT-ELC QRIS in 2015, and what are the prospects for sustainability?	●				
2. What incentives or compensation strategies are most effective in encouraging QRIS participation?	●			●	
3. How effective have Consortia been at fostering an improved early childhood system to support early learning and quality improvement in their region? To what extent have the local QRISs been used to align initiatives and projects at the local level?	●				
QRIS Ratings: Validations and Perceptions in the Field					
4. How effective are the California Common Tiers' structure and components/elements at defining and measuring quality in early learning settings?			●		
5. To what extent do the graduated elements and tiers correspond to graduated increases in child outcomes, including (but not limited to) children's learning, healthy development, social/emotional health, and school readiness?			●		
6. To what extent can the Consortia's local QRIS be streamlined and still result in the same program quality level and same child outcomes? What common elements of the Hybrid Matrix and Pathways are most important to include?			●		

Research Question	System Implemen-tation	Perceptions of Quality and the Ratings	Validity of the Ratings	Quality Improve-ment Supports	Quality and Outcomes
7. In context of the findings of the QRIS descriptive study literature review, are there other tiers, resources, measures, tools, or system structures that should be included that support QRIS reliability, validity, and efficiency in program quality and have led to better overall outcomes in other systems or states?			●		
8. How effective are the Consortia in increasing public awareness of the characteristics of early learning program quality that promote better outcomes for children?	●	●			
Quality Improvement Activities and Changes in Quality & Children's Outcomes					
9. What are early learning staff's experiences with quality improvement activities??				●	
10. How do the QRIS strategies (e.g., technical assistance, quality improvement activities, incentives, compensation, family/public awareness) improve program quality, improve the professionalization and effectiveness of the early learning workforce, and impact child outcomes? Which strategies are the least/most effective?					●
11. For which quality improvement activities does increased dosage (time and intensity of participation) impact program quality and child outcomes?					●
12. What QRIS strategies/variables best impact measurable site progress through the tiers? What barriers exist in progressing though tiers?				●	●
13. What is the cost versus benefit for various QRIS strategies relative to child outcomes?					●

Challenges and Limitations

Several limitations to the study are important to highlight at the outset. First, this is not an experimental study from which causal conclusions about the effects of ratings or QI supports on outcomes can be drawn. Findings are presented in terms of associations rather than impacts. Also, it is important to remember that there may be selection effects that drive observed associations; for example, the most motivated teachers may be the ones who participate in the most QI activities. In addition, the high-quality programs participating in California's QRIS during this initial phase were primarily state or federally contracted programs serving low-income families; the children in the highest quality programs may be more socioeconomically disadvantaged than the children in the lesser quality programs where there are more parents able to pay fees. Thus, results should be used to inform the discussion about RTT-ELC QRIS and its

evolution but should not be used as conclusive evidence to support or undermine specific policy changes.

Second, just over a third of programs across California that are participating in the RTT-ELC QRIS had a full, nonprovisional rating and thus were eligible for inclusion in the study. The study was launched while the RTT-ELC QRIS was still in the early stages of implementation. Many participating programs had not yet received a full rating at the start of the study because they did not have finalized scores on all of the rating elements. The programs with full, nonprovisional ratings differ from programs without full ratings in important ways. For example, they are more likely to receive standards-based funding, such as Title 5 or Head Start, and are therefore already required to meet certain quality standards that other programs are not. This selective group of programs also have limited variability in rating levels, with no Tier 1 sites and few Tier 2 sites.

Full Versus Provisional Ratings

- Programs with full ratings are defined as those with QRIS data that are complete and nonprovisional.
- Complete data are determined by having a QRIS rating and scores on each applicable rating element. (The number of applicable rating elements is determined by the Hybrid Rating Matrix and varies by program type, ranging from four to seven elements.)
- Nonprovisional data further exclude programs awaiting classroom observations and thus without a finalized score on the associated elements.
- For the study, each Consortium identified the sites within its local QRIS that had provisional ratings as of January 2014.
- The study team excluded programs identified as having provisional ratings as well as those without complete data on the QRIS ratings and applicable element scores because inclusion of nonfinalized QRIS ratings would bias the study results.

Third, study samples also were somewhat smaller than the anticipated sample sizes, in part because fewer programs were eligible for the study than anticipated and also because of delays in recruitment in the first year of the study due to extended negotiations with the Consortia. The validity analyses cannot be considered conclusive because the small sample size and lack of variability in ratings among fully rated programs limit our ability to detect differences between each rating level. A further implication of the limited samples is that the study results may not be generalizable to all programs participating in the RTT-ELC QRIS. In addition, the samples of programs that participated in data collection for the validity analyses and outcome analyses included insufficient numbers of FCCHs to permit separate FCCH statistical analysis.

Finally, there are some limitations to the validation research conducted because the RTT-ELC QRIS is relatively new and not fully implemented. As described previously, the state was required to conduct the evaluation within this time frame and validation research that is conducted in the early stages of QRIS implementation can be used to make decisions about revising and improving the system. Although examining the system and how it is performing at this early stage has value and can help the state consider possible revisions to the QRIS, results presented in this report should be interpreted within the context of the system's stage of development and participating programs at the time this evaluation was conducted.

Organization of This Report

This comprehensive report, describing study findings for the Independent Evaluation of California's RTT-ELC QRIS, is organized into eight chapters, including this introductory chapter. Exhibit 1.1 provides a graphical overview of the structure of the report.

Chapter 2 characterizes the status of implementation of the system across the 17 Consortia, including the process of assigning ratings, providing QI supports, and working toward sustainability of the system.

Chapter 3 provides a summary of prior analyses examining comparisons of the tier ratings from the RTT-ELC QRIS against other research-based measures of child care quality—including the CLASS and PQA—to assess the validity of the ratings. This chapter also presents additional analyses that examine the extent to which tier ratings and element scores can be used to predict children's developmental outcomes.

Chapter 4 focuses on the perceptions of quality and the rating system held by parents as well as providers themselves as a way of understanding the potential for the QRIS to influence decision making and practice.

Chapter 5 describes staff-level QI supports—such as coaching, workshops and training, peer support networks, and financial incentives—and their prevalence.

Chapter 6 describes program-level QI supports received—including learning opportunities for directors as well as program-level financial supports.

Chapter 7 examines the extent to which specific QI strategies are associated with increases in early learning program quality and gains in children's learning and development during the course of the program year.

Chapter 8 presents the results of the collection of data from 11 focal Consortia on the cost of QI activities.

Chapter 9 summarizes the key findings presented in chapters 2 through 6 organized by the research questions outlined in exhibit 1.2. We also describe study implications and present some considerations for the future of California's RTT-ELC QRIS.

Chapter 2. Implementation of the RTT-ELC QRIS

California's Race to the Top-Early Learning Challenge Quality Rating and Improvement System (RTT-ELC QRIS) began in 2012. As the federal grant which has supported a pilot of the system nears its end, the QRIS continues to grow and undergo refinement. This report focuses on the 16 counties which have participated in the four-year pilot of the system; however, the remaining 42 counties have now begun to participate in at least the quality improvement (QI) components of the system.

Data and Sample

- Interviews were conducted with the QRIS administrators of all 17 Consortia. See appendix 2A for a list of respondents.

Analysis Approaches

- Interviews solicited feedback on progress that had been made since the interviews conducted in 2014, as well as reflections on the implementation process and plans for sustainability.
- Interview transcripts were analyzed using qualitative data analysis techniques to identify common themes and response patterns.

In order to provide a context for the evaluation findings in this report, this chapter addresses the implementation of California's QRIS, including the status of provider participation, program quality assessments, the Hybrid Rating Matrix, quality improvement initiatives, and the impact of the RTT-ELC QRIS on the alignment of early care and education systems and initiatives. We also report on how the pilot counties have approached the publication of program ratings and how they view the prospects for system sustainability.

The chapter addresses the following research questions:

- What is the status of implementation of the RTT-ELC QRIS, and what are the prospects for sustainability?
- RQ 2. What incentives or compensation strategies are most effective in encouraging QRIS participation?
- RQ 3. How effective have Consortia been at fostering an improved early childhood system to support early learning and quality improvement in their region? To what extent have the local QRISs been used to align initiatives and projects at the local level?

The chapter is informed by interviews conducted in summer 2015 with administrators from the 17 Consortia which served as pilots for the RTT-ELC QRIS system. Our intent was to capture progress that had been made since the first interviews were conducted in 2014. It should be noted, however, that several important developments have occurred since summer 2015. For example, all 58 counties now participate in the First 5 IMPACT Grants, which are designed to support alignment with the California Hybrid Rating Matrix standards in early care and education programs financed by different funding streams. In addition, as of spring 2016, 45 counties participate in the California State Preschool Program (CSPP) QRIS Block Grants and 47 counties in the Infant/Toddler (I/T) QRIS Block Grants. These grant programs provide higher

reimbursements to State Preschool and Title 5-contracted programs meeting higher tiers on the California QRIS matrix.

The following overview summarizes the highlights of the QRIS implementation in the 17 Consortia in the 16 counties participating in the RTT-ELC QRIS pilot as of summer 2015:

- A majority of Consortia had reached their total anticipated number of QRIS participants. As would be expected of a QRIS pilot where participation is completely voluntary, and where resources to support participation are limited, only a minority—20 percent of licensed centers and less than 3 percent of licensed family child care homes in the 16 RTT-ELC counties—chose or were able to participate in the QRIS.

- A majority of the Consortia were on track with program quality assessments. This accomplishment is especially notable because the counties varied substantially in their prior experience with program quality assessments. Although some Consortia had years of experience planning for and conducting valid, reliable and independent classroom Environment Rating Scale (ERS) and Classroom Assessment Scoring System (CLASS) assessments, other Consortia had minimal experience or insufficient resources for these observations, and hence found the work more difficult and costly.

- All Consortia elected to use the common criteria for Tiers 1, 3, and 4 of the Hybrid Rating Matrix. Although RTT-ELC counties had the option to make local modifications in the quality indicators for Tiers 2 and 5, most of the Consortia did not make changes to Tier 2. However, one-third made minor changes to Tier 5, and, at the time we interviewed Consortia leaders, several more counties were considering future changes in one or both tiers.

- In the third year of the pilot (May 2015), at the behest of a few of the Consortia with less experience and resources for program quality assessments, the state revisited the ERS element of the common Tier 3 of the Hybrid Rating Matrix. Consortia leaders voted to modify the ERS element, eliminating the requirement in Tier 3 for a minimum ERS score and allowing a self-assessment. In addition, the Consortia agreed to accept National Association for the Education of Young Children (NAEYC) accreditation in lieu of the ERS score for Tier 5.

- By summer of 2015 all of the Consortia had implemented their proposed quality improvement activities in the areas of workforce, professional development, training and technical assistance activities, with coaching one of the most valued initiatives.

- Most Consortia administrators believed that the RTT-ELC QRIS has enhanced collaboration and alignment among QI activities, promoting a common language among ECE professionals.

- One key feature of a QRIS – publicizing of ratings – had not yet been fully implemented by summer 2015. At that time, the quality ratings resulting from the assessments and other elements of the rating matrix had largely been used internally for purposes of guiding QI activities or determining the level of financial rewards for a participating

program. While the Consortia voiced concerns about publicizing the ratings during a pilot system where participation is voluntary and the rules may change, most planned to make the ratings publicly available by the end of the grant, a goal which, according to the state Implementation Team achieved by the end of 2015.

- Most Consortia thought the California State Preschool Program (CSPP) QRIS Block Grant, the Infant/Toddler (I/T) QRIS Block Grant, and the First 5 California IMPACT grant would together help sustain the system developed during the RTT-ELC QRIS pilot. At the same time, a majority of administrators expect to have to scale back some of their QRIS activities after the federal RTT-ELC funding ends in June 2016. Hence, Consortia leaders are considering various approaches to reducing costs, such as limiting some QI activities to a smaller number of sites, perhaps concentrating on programs in the lower tiers. Others would like to consider further modification or elimination of the Hybrid Rating Matrix requirement to conduct independent ERS assessments.

Overall, we found that the Consortia had implemented the vast majority of the pilot objectives, that they embraced most components of the QRIS system, and that they thought the system helped integrate and improve early learning and care services. Questions remain whether, after the RTT-ELC QRIS federal grant expires, there will be sufficient resources from the new funding streams to motivate participation by a more complete spectrum of private as well as publicly supported providers.

Status of the Implementation of the RTT-ELC QRIS

In this section, we explore in more detail the status of implementation of the RTT-ELC QRIS as of summer 2015 across the 17 Consortia, including reaching provider participation targets and progress toward accomplishing the activities outlined in their plans.

Provider Participation

Two thirds of the Consortia had reached their participation targets by summer 2015; others expect to reach that goal by mid-2016.

As of the time of the interviews, the majority of the Consortia had reached their goals for participation in the QRIS pilot. Several of the Consortia had rolling enrollment, in which more providers joined the system over time, often in what QRIS administrators referred to as “cohorts,” while in at least two Consortia, “participation” is a more appropriate term than enrollment, as sites became part of the QRIS by virtue of being part of the local QI initiative that began before RTT-ELC was implemented. The remaining Consortia expected to reach full participation by the end of December 2015 or mid-2016, and at least three Consortia leaders thought that they would likely enroll more providers than they had projected to enroll in their original proposals to the California Department of Education (CDE).

In 2015, most Consortia had overcome any hurdles to participation that they had faced in previous years; however, some challenges remained. For example, in one Consortium, the administrative agency itself had faced challenges that had impeded full enrollment—staff turnover and diminished capacity had resulted in delays in enrollment in the local QRIS.

The introduction of supplemental funding streams increased enrollment of certain types of programs in some Consortia. For example, according to one Consortium administrator, the CSPP QRIS Block Grants "completely changed the landscape for school district enrollment." Although some State Preschool sites in this Consortium had been reluctant to join the QRIS as of 2014, in 2015, the incentive for the reward money associated with the CSPP QRIS Block Grants overcame their resistance.

Enrollment would need to be significantly expanded to represent all types of programs.

Several Consortia reported facing challenges recruiting providers in particular settings, such as family child care homes (FCCHs) or private programs supported by parent fees. The publicly funded and especially publicly contracted programs in many Consortia have a long history of participation in QI initiatives, and there have frequently been financial and other incentives for their participation in these activities. The privately funded programs have less experience with publicly administered QI initiatives and, short of more financial or other incentives, may be more wary of participation.

As proposed by the California Early Learning Quality Improvement System (CAELQIS) Advisory Committee in 2010, the QRIS system was never expected to include all licensed programs initially. Rather, the proposal was to begin with a pilot lasting at least three years, followed by phased-in implementation over five years or more. Moreover, the vision was that participation in the QRIS would initially be voluntary, then be required for publicly funded programs, and ultimately be required for all licensed programs with appropriate funding and incentives provided (CAELQIS 2010).

Even with the receipt of the RTT-ELC supplemental funds in 2013, there were not sufficient resources to provide assessments, much less incentives, for all providers to participate. Moreover, the federal RTT-ELC grant requirements emphasized activities focusing on programs serving children with high needs, i.e., children from low-income and otherwise disadvantaged populations. By 2014, 1,272 programs were participating in the 17 RTT-ELC Consortia in the 16 counties, representing 4 percent of the total number of licensed settings in the 17 Consortia (n=30,271) and 3 percent of the total number of licensed settings in the state (n=41,931) (exhibits 2.1 and 2.2). By September 2015, according to the online QRIS Compendium, enrollment had expanded to 2,232 sites, representing 7 percent of the total number of licensed settings in the 17 Consortia and 5 percent of the total licensed settings in the state. Most of the licensed sites participating are centers, with very few family child care homes in the system. Thus, by 2015, participation of centers reflected 20 percent of the licensed centers in the 16 RTT-ELC counties and 15 percent of licensed centers in the state. A majority of the participating centers were state or federally contracted programs, such as State Preschool or Head Start.

Many Consortia leaders recognize that a QRIS will only achieve its full potential when all program types participate. As one administrator noted, the real value in a QRIS is when the various programs in a community can be fully engaged and rated, thereby creating the market pressure for providers to participate and providing valuable information on the full array of ECE choices to parents. Given the federal focus on using the RTT-ELC funds to reach out to programs serving disadvantaged children, the voluntary nature of the QRIS, and the limited incentives to offer privately funded programs, the narrow segment of providers participating in

the pilot system would be expected. However, as will be discussed in future chapters, the lack of broad representation of all types of providers does restrict the interpretation of the evaluation findings; had a more representative sample of providers participated, the findings might have been different.

Exhibit 2.1. Number of Licensed Settings in the 17 Consortia and the State, as Compared With Programs in the QRIS, 2014

	Number of Licensed Sites in the County*			Number of Sites in the QRIS**		
County	**Centers**	**FCCHs**	**Total Sites**	**Centers**	**FCCHs**	**Total Sites**
Alameda	568	1,502	2,070	12	5	17
Contra Costa	355	990	1,345	41	21	62
El Dorado	62	94	156	24	8	32
Fresno	290	634	924	36	14	50
Los Angeles	2,783	7,378	10,161	245	74	319
Merced	71	229	300	36	12	48
Orange	846	1,301	2,147	65	3	68
Sacramento	466	1,445	1,911	103	30	133
San Diego	960	3,693	4,653	86	15	101
San Francisco	311	697	1,008	97	14	111
San Joaquin	181	612	793	26	47	73
Santa Barbara	174	363	537	69	28	97
Santa Clara	666	1,867	2,533	17	2	19
Santa Cruz	114	332	446	40	0	40
Ventura	240	738	978	51	24	75
Yolo	86	223	309	27	0	27
Total in 16 counties in QRIS	**8,173**	**22,098**	**30,271**	**975**	**297**	**1,272**
Total in State of California	**11,230**	**30,701**	**41,931**	**975**	**297**	**1,272**

*SOURCE: California Child Care Resource and Referral Network's *2015 Child Care Portfolio*

**SOURCE: 2014 Common Data Elements, as cited in AIR and RAND (2015).

Exhibit 2.2. Number and Percentage of Licensed Sites Participating in the QRIS in 2014 and 2015

	Number of Sites in the QRIS in 2014*			Number of Sites in the QRIS in 2015**		
	Centers	FCCHs	Total Sites	Centers	FCCHs	Total Sites
Total number in 16 counties	975	297	1,272	1,652	580	2,232
Percentage of licensed sites in 16 counties	11.9%	1.3%	4.2%	20.2%	2.6%	7.4%
Percentage of licensed sites statewide* **	8.7%	1.0%	3.0%	14.7%	1.9%	5.3%

*SOURCE: 2014 Common Data Elements, as cited in AIR and RAND (2015).

**SOURCE: QRIS Online Compendium, 2015

***SOURCE: Percentages calculated using the total number of licensed sites reported in Exhibit 2.1 from the California Child Care Resource and Referral Network's 2015 Child Care Portfolio.

Program Quality Assessments

Despite challenges, the Consortia have managed to implement program quality assessments in accord with the *Consortia Implementation Guide for the Tiered Quality Rating and Improvement System (TQRIS*) (California Department of Education 2015). To address the costs of the assessments, the Consortia have continued to look for creative ways to improve the efficiency of their administration; and the state Implementation Team composed of CDE and First 5 California has assisted the Consortia in seeking solutions. RTT-ELC funds were used to contract with First 5 California to establish an inter-rater reliability system, and with the assistance of the Assessor Management Workgroup, has developed an assessor management system.

Administration of Assessments

To implement the requirement for independent assessments, most Consortia relied on external, contracted staff, although the specific responsibilities of assessors varied.

Although two of the Consortia used both internal and external contracted staff for conducting ERS and Classroom Assessment Scoring System (CLASS) assessments, the majority of the Consortia relied solely upon external, contracted staff for the work. Among these Consortia, the responsibilities that assessors assumed varied. For example, in one Consortium, assessors were responsible for scheduling, conducting the observation itself, compiling a summary report, entering the collected data, fielding questions from the administrative agency, and fielding questions from the sites and/or site coaches, as needed. In other Consortia, assessors' roles were more limited and the administrative agency assumed some of the labor. In one Consortium, for example, administrators noted that although observers were responsible for conducting the observations, entering data and submitting data reports, and answering any clarifying questions from the administrative agency, the cost of the assessments did not include scheduling.

The recent CSPP QRIS Block Grant and I/T QRIS Block Grant funding directly link incentive dollars to ratings. This linkage may increase the stakes for assessments. Because of that change,

as one QRIS administrator stated, the administrative agency needs "ironclad confidence in their assessment team."

Consortia followed the Consortia Implementation Guide for the frequency of and selection of classrooms for assessments, with a few Consortia exceeding the minimum requirements.

At a minimum, the *Consortia Implementation Guide* requires Consortia to assess centers and FCCHs every other year with the age-appropriate version of the CLASS and the age- and setting-appropriate version of the ERS. As of summer 2015, all but one of the Consortia stuck to these minimum guidelines for observations. The one Consortium that conducted observations more frequently than required, as in 2014, continued to conduct both CLASS and ERS observations on an annual basis. However, in the 2015–16 program year, programs rated at "4" or "5" could "lock in" their rating for two years (that is, elect to not be rated on an annual basis). After 2015–16, this same Consortium would assess all of its programs every other year. In addition, in two Consortia, assessors observed every session, not just the sample that is required by RTT-ELC. Finally, a few Consortia leaders mentioned having to assess more frequently than anticipated at some sites because reassessments were required when there is significant change to the site (for example, teacher turnover at a small site).

In a few Consortia, some of the providers in the QRIS pilot requested additional time before their assessments were conducted in order to make necessary quality improvements and, in some cases, familiarize themselves with various assessment tools (such as CLASS and ERS) that were new to them. These providers thought that extra time would allow them to make the necessary program enhancements that could improve their rating. But when the CSPP QRIS Block Grant was introduced, State Preschool program staff in a few Consortia wanted to be assessed more quickly than usual so that they could determine whether they were a Tier 4 or 5 and therefore be eligible for a local CSPP QRIS Block Grant.

Affordability, Manageability, and Sustainability of Program Quality Assessments

While many Consortia reported challenges finding, training, and retaining qualified classroom assessors in 2014, most of these challenges had lessened by summer 2015.

In 2014, challenges associated with classroom observations included cost and logistics, such as scheduling. The most commonly cited challenge in 2014 related to finding or keeping a sufficient number of assessors. Several counties did not have a cadre of trained and reliable assessors to draw from when needed and thus had to train a pool of people. By 2015, although a couple of leaders noted that it was still somewhat more challenging to find assessors for the infant/toddler-age assessment tools than it was for the preschool-age tools, most Consortia had overcome the challenges of finding assessors.

As in 2014, the cost of observations varied considerably among Consortia in 2015, though average costs of observations remained high.

Four counties in one region had decided upon a similar rate to pay their CLASS and ERS assessors; however, most Consortia had determined their own rates. In 2015, the cost of the CLASS observations ranged from $300 to more than $1,000 per observation. Similarly, among

those Consortia that reported on the costs of their ERS observations, the rates also ranged from $300 to more than $1,000. But, as in 2014, the ERS observations, on average, continued to be more expensive than the CLASS observations. The services included in assessment costs varied by Consortium, thereby accounting for some of cost differences. For example, in some Consortia, assessors' roles were limited to conducting the assessment, entering data, and submitting data reports, while in other Consortia, assessors were responsible for all aspects of the data collection and reporting, from scheduling to reporting and fielding questions from the sites and/or site coaches, as needed.

Decisions to change assessor teams within Consortia as well as modifications to the Hybrid Rating Matrix are expected to increase the efficiency of program quality assessments.

In a few Consortia, administrators had changed or were planning to change assessors to develop local capacity and/or reduce costs, thereby fostering affordability and sustainability. For example, in one Consortium, the administrative agency had switched contractors, and the new contractor had assumed more responsibility. In another Consortium, administrators had issued a request for proposal in summer 2015 for a group to conduct CLASS and ERS assessments in an effort to contract with and develop their own pool of assessors in "order to manage the process more closely and efficiently." In a third Consortium, administrators were trying to build local capacity with another agency that is local and charges a bit less than their current contractor.

One significant change cited by several leaders that would impact assessment costs was the change to the ERS Element of the Hybrid Rating Matrix at points three and five in mid-May 2015. One leader indicated that the ERS assessments would become a sort of a "tipping point" rather than a requirement, which would reduce costs considerably. In other words, they would determine whether an ERS assessment was necessary based on the number of points that a site had earned from the other six elements in the Hybrid Rating Matrix. For example, if a site were just one point away from the next tier on the rating matrix, the administrative agency would devote resources to having an ERS assessment done at the site. But the agency would not conduct an assessment if a higher ERS score would not earn the site enough points to push it to the next tier.

The inter-rater reliability contract and the assessor management system developed at the state level are intended to help support the quality of the assessments.

First 5 California hired three individuals who became certified by the ERS Institute (ERSI) at a minimum of 90 percent reliability to serve as State Master Anchors and who also are certified by Teachstone as Affiliate CLASS Trainers. The State Master Anchors also certify local/regional anchors at no cost to the Consortia, a cost savings of about $6,000 per anchor had they gone to ERSI for the same service. CLASS training is also provided several times a year with calibration three times per year on each CLASS tool. The RTT-ELC QRIS Assessor Management work group developed an RTT-ELC Assessor Management Structure to address the roles, responsibilities, and relationships among the State Master Anchors, local Anchors and local assessors for the ERS and CLASS tools. This work is intended to ensure ongoing quality control through the development of reliability/calibration standards for CLASS and ERS.

The Hybrid Rating Matrix

As indicated above, the implementation of the Hybrid Rating Matrix varied somewhat at the local level, thereby limiting to a degree the comparability of the ratings across Consortia and providing important context for understanding the validation analyses included in this evaluation. For example, Consortia also had the option to block Tier 2 and make modifications to Tier 5 of the Hybrid Rating Matrix, and although most of the Consortia were maintaining the common criteria for Tier 2, five Consortia had made changes to Tier 5. Some Consortia also had considered making changes to Tier 2 or 5 in the future.

Local Adaptations to Tiers 2 and 5

The vast majority of Consortia were maintaining the common criteria for Tier 2, although several were considering future changes to this level.

As of the time of the interviews in 2015, 14 of the 17 Consortia were maintaining the common criteria for Tier 2. The most common reasons for not modifying Tier 2 were to avoid complicating the process, to ensure consistency among Consortia, and to prevent the addition of criteria that would create more burden for early care and learning staff. One Consortium noted that changing Tier 2 was irrelevant because it did not have any Tier 1 or 2 sites—it only had sites in Tiers 3, 4, and 5.

Only two of the 17 Consortia had made changes to Tier 2, electing to change Tier 2 from a point structure to a blocked structure (as is done for Common Tier 1). An additional Consortium noted that they added the requirement to be familiar with two documents, *California Preschool Learning Foundations (Foundations)* and *California Preschool Curriculum Frameworks (Frameworks).*

In three Consortia, administrators were considering making modifications to Tier 2 in the future, most likely after the conclusion of the RTT-ELC grant. In one of these three Consortia, a consultant who has been on board since the beginning of the grant was helping them evaluate local changes (for example, changing Tier 2 from a point structure to a blocked structure). In two different Consortia, leaders were considering adding requirements related to fostering inclusion; one of these administrators also discussed adding requirements related to family engagement and cultural sensitivity. One Consortium was thinking of revisiting and reconsidering the director qualifications in the rating matrix.

Nearly a third of Consortia either had made modifications to Tier 5 or were considering making adjustments to it in the future.

As of summer 2015, in 12 of the 17 Consortia, administrators were maintaining the criteria to receive five points on each element as delineated in the Hybrid Rating Matrix. As with Tier 2, the most common reasons for not modifying Tier 5 were to ensure consistency across Consortia and reduce burden on providers. As was the case in 2014, two Consortia had made changes to both Tier 2 and Tier 5.

Five Consortia had modified Tier 5. In two of these Consortia, leaders had added professional development requirements for directors at Tier 5. Other modifications included requiring professional development hours related to serving English learners and children with special needs, additional requirements for assistant teachers, requiring teacher training on the *Foundations* and *Frameworks*, requirements intended to strengthen families, and required accreditation.

Four additional QRIS administrators were considering making changes to Tier 5 in the future. Changes under consideration included adding an inclusion requirement, including requirements concerning director qualifications, adding requirements related to early literacy, and emphasizing that sites use evidence-based curricula (sites could select their own) and lesson planning. Whether or not leaders had made changes to Tier 5 as of summer 2015, a few administrators worried more broadly about how changes might be made and about how they might impact consistency among Consortia and future work. For example, one leader said, "When and how and with what kind of support will we tackle modifications to the matrix? It's the big elephant [in the room]."

Modifications to the ERS Element at the Three-Point Level and the Five-Point Level

During a Consortia meeting held in May 2015, attendees voted to change the three-point level of the common, five-tiered Quality Continuum Framework Hybrid Rating Matrix to eliminate the requirement for a minimum ERS score and allow a self-assessment; the five-point level was changed to accept National Association for the Education of Young Children (NAEYC) accreditation in lieu of the ERS score requirement. The changes to the ERS element of the rating matrix are shown in italics in exhibit 2.3.

Although the May 2015 decision to modify the ERS element at the three-point and five-point levels was based on a consensus vote, Consortia leaders' perspectives varied on the topic. During the interviews, some QRIS administrators were firm proponents of the change, while others had mixed feelings, and some expressed serious concerns about the decision, as described in further detail below.

Exhibit 2.3. Changes to ERS Element (Program Environment Rating Scale) of the Quality Continuum Framework Rating Matrix at Points 3 and 5

Element	Time Period	Block	2 Points	3 Points	4 Points	5 Points
Program Environment Rating Scale(s) (Use tool for appropriate setting: ECERS-R, ITERS-R, FCCERS-R)	Pre-May 2015	☐ Not required	☐ Familiarity with ERS and every classroom uses ERS as a part of a quality improvement plan	☐ Independent ERS assessment using scale for the appropriate setting; All subscales completed and averaged to meet overall score level of 4.0	☐ Independent ERS assessment. All subscales completed and averaged to meet overall score level of 5.0.	☐ Independent ERS assessment using scale for the appropriate setting; All subscales completed and averaged to meet overall score level of 5.5.
	Post-May 2015	☐ Not required	☐ Familiarity with ERS and every classroom uses ERS as a part of a quality improvement plan	☐ *Self-* assessment on the whole tool. Results used to inform the program's quality improvement plan.	☐ Independent ERS assessment. All subscales completed and averaged to meet overall score level of 5.0.	☐ Independent ERS assessment. All subscales completed and averaged to meet overall score level of 5.5 OR Current national accreditation approved by the California Department of Education

NOTE: ECERS-R= Early Childhood Environment Rating Scale-Revised; ITERS-R= Infant/Toddler Environment Rating Scale-Revised; FCCERS-R= Family Child Care Environment Rating Scale-Revised

Changes at the Three-Point Level

Consortia identified reduced administrative costs due to decreased demand for independent ERS observations as a potential benefit of the change to the three-point level on the ERS element.

Several Consortia administrators were strong advocates for the change to the three-point level of the ERS element, primarily because it would reduce the number of sites that needed to be assessed with ERS and therefore lower costs. For example, in one Consortium, an administrator noted that more than 60 percent of the expenses for one of its cohorts had been devoted to the

ERS assessment, resources that could have been used to provide resources, professional development, and coaching to the sites instead. In another Consortium, the change would reduce administrative costs because the ERS assessment would only be conducted in sites where there was a possibility that a change in the ERS score would raise a site's rating to a higher level.

Even among those Consortia leaders who did not agree with the change to the three-point level, some said they understood that the change could result in cost savings, particularly for smaller counties and/or counties with fewer years of experience conducting ERS assessments. For example, one QRIS administrator explained that although it would not be a cost savings for her own Consortium, she "voted for that [change] to support the counties that were able to make that work as a cost savings."

The change in the three-point level of the ERS element might also help relieve the burden on individual classrooms and sites as well. For example, one QRIS administrator noted that some of her sites have been overwhelmed with the independent observations that are required by different programs, adding, "Right now, some sites have said that they would rather not have [another] observation or have too many [assessors] come through." As a result, these sites were rated at the two-point level – not because they had received a low score on an independent ERS assessment, but because they had not received an assessment at all. With the change at the three-point level, however, these sites would now be able to use self-assessments or external assessments conducted for other programs (as long as they inform a quality improvement plan) and therefore earn a '3' on the ERS element.

QRIS administrators expressed differing opinions on the use and benefits of an ERS self-assessment.

Several Consortia leaders argued that their sites benefit from the process of conducting an ERS self-assessment. For example, one administrator who was a proponent of the change at the three-point level said that the value of a site's self-assessment was to serve as a baseline to help determine the needs and/or QI services that each of their sites receives before having an independent assessment.

In contrast, a few QRIS administrators advocated for independent ERS assessments as a means of providing an objective baseline measurement. As one administrator noted, "Our experience really consistently has been that the transition from self-assessment to a calibrated, external assessor is a really, really tough one...What happens is the calibrated, reliable external assessor's scores are lower." Sites that have conducted a self-assessment might be surprised and/or upset by an external rater's assessment. From this perspective, for purposes of baseline, it is important that providers receive valid, reliable scores from the outset.

Other concerns expressed about allowing an ERS self-assessment at Tier 3 were that without external pressure on the program administration to make changes to the environment and the appropriate resources to support those changes, necessary improvements would likely not happen. In the absence of a well-developed and designed training and technical assistance system that informs a quality improvement plan, as one Consortium leader noted, a self-assessment does not always change practices much; hence, the Consortium was hesitant to agree that a self-assessment deserves three points.

The change to the ERS element may eliminate one of the main benefits of independent ERS assessments, namely their use for sites in greatest need of improvement.

Another concern raised by two administrators had to do with using an external assessor only in the determination of whether a site is at a four-point or five-point level on the ERS element, as opposed to providing an objective starting point for lower-tier programs to improve. As one administrator explained:

> "Our position is that using it only at [Points] 4 and 5 makes no sense. I care not at all about the score differential on the ERS between [Points] 4 and 5. I don't think that's where you make a difference. I care a lot about a 'sub-2' ERS [score] and getting that up to a four. The ERS isn't the place where you're going to tweak a high quality site…The strength of the ERS… is in helping sites think about some of the environmental and health and safety basics. And so, the way this is being done with only [points] 4 and 5 just doesn't make any sense to me because where the ERS is a helpful tool in our experience is really helping sites that are struggling with environment to come up to the minimum. Not turn the corner on high quality."

Eliminating the minimum score requirement to receive three points on this element is parallel to the CLASS element where minimum scores are required for Points 4 and 5 but not Point 3.

Consortia also raised concerns about the clarity of the change and how the change affects equity and validity.

Other concerns about changing the three-point level related to validity of the rating matrix across sites within and across Consortia. As one administrator explained, "We felt very strongly [that] it wasn't fair to have some [sites] that were rated one way and some that were rated another; it doesn't help with validity." Another administrator noted that the guidance was not "articulated very well in the implementation guide yet, which is really more where the details come out…I don't think they've really clarified for people going out and doing ratings, how to actually score that 3-point level."

Changes at the Five-Point Level

In regard to changes at the five-point level of the ERS element, administrators' opinions were mixed; however, this change seemed less contentious than changes at the three-point level.

Consortia administrators highlighted the potential benefits of the change to the five-point level on the ERS element, such as reducing ERS-related costs and aligning the QRIS with another system of quality rating.

The change at the five-point level, like that at the three-point level, may reduce costs. That is, Consortia with multiple NAEYC-accredited sites participating in the QRIS would have fewer sites that needed independent ERS assessments (because their accreditation would mean that they automatically qualified for five points on the ERS element). As one administrator noted, the change to the five-point level of the ERS Element acknowledges and rewards the hard work that goes into accreditation.

Another perceived potential benefit was the opportunity to promote alignment among systems and definitions of quality. As one administrator explained, “It’s important that we look at how we can align more and more going forward because it’s not good to have different systems of definition of quality.”

Other Consortia disagreed with the change, stating that the ERS element and NAEYC accreditation measure different aspects of quality, but more than a third felt the change would likely not impact them at all.

A few leaders had stronger reservations about the change at the five-point level than the changes at the three-point level because they think NAEYC accreditation measures different aspects than does the ERS. As leaders from one Consortium explained, “I totally understand some of the criticisms of this decision, because it’s not like an NAEYC accreditation focuses on environments the way that the ERS scales do. They are not interchangeable.”

More than a third of Consortia administrators reported that the change would likely not impact them much—or at all—as there were so few NAEYC-accredited centers in their counties. As one administrator explained, there was only one NAEYC-accredited site in her county, and that site indicated that, due to the expense of the accreditation process, the site would prefer to get rated with ERS when its accreditation expired. A second administrator noted that although her Consortium used to have many accredited centers, “a lot [of the centers] hadn’t ‘kept [their accreditation] up.’” Another indicated that most of her Consortium’s accredited programs were FCCHs, not centers, and the change in the rating matrix did not include family child care accreditation. In the future, the state may accept other accreditations (after review of the accreditation criteria) in lieu of a completed ERS assessment as well.

Workforce, Professional Development, Training, and Technical Assistance

QRIS administrators also were asked to share feedback on the status of QI activities, whether any changes had been made since the 2014 interviews, and which activities they felt were particularly successful in improving quality.

Consortia reported, in 2015, offering a range of QI activities, including coaching, often tailored to the needs of the sites.

QRIS administrators noted that all of their proposed QI activities were under way as of summer 2015. Consortia reported providing various QI supports, including coaching, training, technical assistance, and workshops. Several administrators noted that coaching was an integral part of their QI system and described various types of coaching, which were provided to sites based on their baseline rating. For example, in one Consortium, sites at Tiers 2 and 3 have “generalist coaches” who provide coaching on a variety of topics. In another Consortium, coaches provide “informal ratings” before formal ratings are done, so that the coaches can determine necessary supports. A third Consortium makes coaching services available at every classroom session that is part of the RTT-ELC QRIS.

Administrators also indicated that they offer a range of training and technical assistance (TA) opportunities, including CLASS or Desired Results Developmental Profile (DRDP) training or TA services, such as mental health consultation. Some Consortia also offered college coursework to their participants. For example, in one Consortium, administrators contracted with a local community college to offer classes that were held in the community or on college campuses, at times of the day and week that made it easier for the participants to attend.

According to administrators, coaching was one of the most valued QI activities.

Several QRIS administrators listed coaching as one of the most effective QI supports. Even sites rated at Tier 4 or above wanted to continue coaching to help maintain their high rating, as one administrator noted. Not only does the sites' eagerness to continue the coaching "speak volumes to [the wisdom of] our [QRIS] design," this administrator added, but also it is important to recognize that the sites would never be able to afford to purchase this kind of technical assistance on their own.

The coaching TA (Technical Assistance) not only benefits a Consortium's cadre of coaches, according to another leader, but also contributes to the overall administration of the QRIS, adding, "That's how we get our assessors and our site improvement plan… that's huge to us." Another administrator shared, "I believe the best way to ensure that things are well implemented is through a coaching model."

Many administrators said that they had made minor changes or additions and/or refined their system of QI supports since initial implementation.

Nearly a third of administrators added or modified some QI activities over the past year, particularly related to training. Administrators from two different Consortia, for example, increased their CSEFEL (Center on the Social and Emotional Foundations for Early Learning) training capacity. Another Consortium increased training efforts related to the DRDP over the past year, sending several leaders to the trainer of trainers institute for DRDP to increase local training capacity. Leaders in this same Consortium also were considering myTeachstone,[4] an online subscription service for coaches and early learning staff with resources and an array of videos that highlight examples of the 10 dimensions of the CLASS tool, which would support coaching on CLASS scores. Another Consortium added a family child care education project that aims to help family child care providers take up to six units in a year and earn their Child Development Assistant Permit.

Other Consortia made changes to their coaching approach, providing more coaching than in 2014 or having coaching include additional professional development training. Or leaders adjusted the coaching, providing sites with a menu of choices (for example, in-person coaching on CLASS or ERS, My Teaching Partner) rather than providing the same coaching on the CLASS tool to everyone.

---

[4] http://info.teachstone.com/video/myteachstone

Several administrators noted that they refined the QI activities as they went along. For example, one Consortium leader shared that in the initial years, administrators had "brought [in] every training on the matrix available" but in the past year, they had been much "more thoughtful and intentional about which trainings that we bring forward so that we can maximize our time and build capacity around not only their [sites'] interests, but where we think we're getting the most bang for our buck."

Some administrators reported leveraging resources from other funding streams and initiatives to provide QI activities and supports to sites.

In a few Consortia, some RTT-ELC QI supports were leveraged with other funding streams and initiatives. For example, in one Consortium, sites could use AB 212 funds to participate in CLASS training. In another Consortium, administrative staff collaborated with on-site coaches where possible. For example, if a Head Start site had ECE specialists (such as coaches) on staff, the leaders would try to bring those individuals on board to help co-facilitate training efforts. Another Consortium added a professional learning community project for its RTT-ELC participants who are funded by State Preschool. Through this project, teams that include lead teachers and assistants analyze DRDP data on an ongoing basis, and team members receive stipends to attend trainings and extra pay if needed to attend bimonthly professional learning community meetings.

Enhancement and Alignment of Early Care and Education Systems and Initiatives

As they reflect on the RTT-ELC QRIS pilot, Consortia leaders generally think the system has enhanced collaboration with other QI initiatives. Many also think that the two new funding streams—the First 5 IMPACT Grants and the CSPP QRIS and the I/T QRIS Block Grants will help continue the trend toward alignment with the QRIS.

The RTT-ELC effort has enhanced collaboration and alignment among QI initiatives and programs in its counties.

QRIS administrators stressed that the RTT-ELC QRIS has enhanced the ECE system, including helping to improve alignment among its QI systems, particularly regarding communication and collaboration among agencies and systems. For example, one Consortium has established a coaching collaborative, through which it convenes all of the ECE coaches from the various programs and initiatives (such as Child Signature Program and Head Start) across the county on a monthly basis to share resources, discuss each other's coaching practices, and provide some training and support in order to improve their coaching skills. In situations where there was overlap between coaching services, the Consortium and the coaching collaborative were able to work with the administrators and make some clear decisions on how best to utilize the coaches. According to another leader, the RTT-ELC QRIS has enabled the county to reach out to sites and programs that never before would have participated in quality improvement efforts.

The RTT-ELC effort has helped promote a common language among ECE professionals within and across programs and agencies.

Similarly, Consortia pointed out that the RTT-ELC grant has supported collaboration and the use of a common language among administrators and line staff across settings and agencies. As one administrator noted, "I think, even at the federal level, with our Head Start Program and the Early Head Start Child Care Partnership, we're all talking the same language. Whether it's First 5 California or … the state department or the feds at the Head Start Program, we're all on the same page."

As one important indication of the degree of collaboration promoted by the RTT-ELC QRIS, LAUP and LA County Office of Education's STEP program, which had been separate Consortia within the same county, have agreed to merge. Prior to the RTT-ELC QRIS pilot, these two approaches to a QRIS were very different but now they are merged into one. Doing so amounts to a reduction in duplication of effort and is, according the state Implementation Team, a credit to the local control model.

While all counties are participating in the new First 5 IMPACT Grant and the vast majority in CSPP and Infant/Toddler QRIS Block Grant Programs, a few administrators voiced concerns that they would miss the coherence of one dedicated funding stream for QRIS activities, namely the RTT-ELC grant.

All counties were eligible to apply for Phase I of the First 5 IMPACT grant in July 2015, and, as of this spring 2016 writing, all 58 counties have received the grant, with 39 participating in Step 3 which requires that a majority of settings receiving grant funds receive ratings and 19 in the Step 2 phase which requires that settings participate in QI activities but does not require all to receive ratings. Forty-five of the counties also receive CSPP QRIS Block Grant funds, which requires that all sites (essentially the Title 5 State Preschool sites) be rated. For those counties that were not CSPP QRIS Block Grant recipients as of the application due date, they will be able use First 5 IMPACT funds from Year 1 to develop a QRIS and conduct initial ratings of CSPPs. Once CSPP QRIS Block Grant funds have been secured, those funds must support the rating and quality awards for sites with CSPP funding. The First 5 IMPACT grant will provide a total of $190 million over five years.

It will be important to watch the extent to which the new grants will help counties recruit private and non-State Preschool and Head Start programs to participate in the QRIS. A few QRIS administrators shared concerns about potential "siloing" of programs. For example, the CSPP QRIS Block Grant targets state-contracted State Preschool programs, providing higher reimbursements for them to reach Tiers 4 and 5 on the RTT-ELC Hybrid Rating Matrix. Centers that provide vouchers under the Alternative Payment Program or CalWORKs may participate only if they also have CSPP contracts. A broader spectrum of providers, such as private providers and family child care homes, is eligible to receive assistance under the First 5 IMPACT grants. However, the assistance is limited primarily to QI activities, such as program quality assessments and coaching, and typically does not include the incentive of higher program reimbursements."

Next Steps

Public Awareness

QRIS administrators were asked to explain when and how they planned to publicize providers' ratings, and early learning staff were asked to provide feedback about the ratings and the Hybrid Rating Matrix that informed them. As of summer 2015, Consortia had primarily used the ratings for internal purposes, such as to inform the type of QI needed by a site, and, in some cases, to determine the level of payment or reward for which a site qualifies. However, since that time, according to the state Implementation Team, all of the pilot counties have published ratings.

Planned Rollout for Publicizing Program Ratings

As of summer 2015, one of the 17 Consortia had publicized ratings, with most planning to do so by the end of the year.

The RTT-ELC pilot funding required Consortia to make their ratings available to the public by the end of the grant. As of August 2015, one Consortium had published access to a searchable, online database hosted on the website of the community's Resource and Referral (R&R) agency, which includes those programs that have scored a "3," "4," or "5" per the Rating Matrix.

Most Consortia reported that they were planning to release ratings in either late 2015 or spring 2016. The majority of these Consortia leaders shared that they expected the ratings and additional information about quality to be posted by the end of 2015. One leader noted that because all of the ratings for all of its cohorts would be done in spring 2016, administrators anticipated making them available online around that same time period.

A few administrators were considering aggregating ratings by area and setting and sharing more general information about the quality of programs in their county. For example, one leader said that her Consortium might present the percentage of family child care providers in the western part of the county at Tier 1, 2, 3, 4, or 5.

Although most of the Consortia had not yet publicized their ratings, information about involvement in the QRIS was shared in other ways. For example, a few Consortia had publicized the list of providers in the system as of August 2015. As one leader noted, "[W]e are publicizing everyone who is participating in a QRIS as already going above and beyond to participate in continuous quality improvement." In addition, providers in some Consortia were already making their ratings available to the public on their own.

Consortia used a collaborative approach as they worked to develop their plans for sharing ratings and information about quality.

Administrators from several Consortia said that although they had not yet made their ratings publicly available, they were working with their partners to develop a plan to share ratings and other information about quality. For example, one leader noted that a work group and communications specialist were devoted to developing a plan for publicizing ratings and an accompanying website. In another Consortium, as of summer 2015, administrators noted that

they were finalizing their RFA process for a public awareness campaign about QRIS. Another leader noted that the local planning council had an advisory committee for QRIS implementation that is spearheading the conversation about publicizing ratings.

In several Consortia, the administrative agency was soliciting input from various agencies in the community, including representatives from R&R agencies, the local child care planning agency, four-year colleges, community colleges, school districts, and special education services. Administrators also reported seeking advice on disseminating information about quality from their partner county offices and First 5s in other counties within the region.

Consortia discussed how best to use the R&R to help with dissemination and promotion of the QRIS. One administrator reported deliberating on whether R&Rs would just share whether a site participated in the QRIS or whether they would share the rating itself. In another Consortium, leaders said that they have asked their R&R to prioritize enrollment in the local quality initiative when they make a referral.

In making plans to share the ratings, Consortia focused on how to convey the information clearly to parents and support their understanding of quality.

As of summer 2015, much of the planning had focused on how to explain quality to parents and other stakeholders. For example, the strategy in one Consortium did not include specific plans for highlighting any particular programs, but provided more general information about the different elements in the rating matrix and what constitutes quality early learning and care. In fact, many QRIS administrators noted that ratings should not be shared unless they are accompanied by an explanation of what the individual components (and quality in ECE) mean in layperson's terms. As one leader stated:

> "In this community, we have focused more on trying to describe what it means to be a tier three or four, rather than saying this program is a tier three or four, similar to an individual who gets a grade report. If someone gives you a 'B' on something, what does that 'B' really mean? And so we're trying to use more descriptive language to describe what the rating means as opposed to actually tagging a site with a number."

In another Consortium, a leader noted that the plan was to make the ratings as easily understandable and as positive as possible and strive to communicate quality along a continuum without using a number or stars. As such, this Consortium grouped the five tiers into categories: (1) Emerging Quality, or "Bronze" (Tiers 1 and 2); (2) Quality, or "Silver" (Tier 3); and (3) High Quality/High Quality Plus, or "Gold" (Tiers 4 and 5). A leader from another Consortium noted that the Consortium's work group was trying to decide between posting all five tiers or aggregating sites into groups (for example, "1" and "2" might be "silver," "3" might be "gold," and "4" and "5" might be "platinum"). This approach attempts to assign familiar terms to help parents understand the ratings at a broader level, although having the same term hold different meanings in different Consortia (for example, "silver" means "3" in one Consortium but "1" and "2" in another) may create some confusion across counties.

In at least one Consortium, administrators were considering convening parent focus groups to solicit input from families about ratings. A leader from this Consortium noted that administrators

wanted parents to have sufficient information, but not so much information that they were overwhelmed.

Some Consortia leaders reported that they planned to share other information beyond just the ratings themselves. For example, one Consortium leader noted that the goal of publishing ratings is to help parents understand what is meant by quality and to explain what a rating means instead of just providing a numerical score. This administrator added, “So, we are being very careful as to how we're going to message that to the public.”

While engaged in activities to plan for publicizing the ratings, the majority of Consortia leaders expressed some reservations about doing so during the pilot phase of the QRIS.

QRIS administrators expressed reservations about publicizing ratings during the pilot. The primary concerns were as follows: (1) the ratings were based on a pilot program in which participation was voluntary and did not reflect the full spectrum of early care and education settings; (2) the rating system is still a work in progress and the rating matrix itself is likely to change; (3) there is not sufficient authority to require, or resources to attract voluntarily, private programs to participate; (4) the ratings may not have been updated frequently enough during the pilot to be fair to sites; and (5) compared to the broader spectrum of private fee-paying or publicly supported voucher families, the ratings may be of less use to parents of children in state- and federally contracted programs (the majority of the programs represented in the pilot)—who have limited choices in selecting the child care program for their child and have enrolled in these programs because of financial need, rather than other factors.

First, several QRIS administrators had reservations about publicizing ratings that were based on pilot program that relied on providers to volunteer to participate. A few leaders noted that they had encouraged provider involvement by emphasizing that the QRIS was a pilot and assuring recruits that it was a “low stakes” initiative representing little risk to the provider. This sense that the rating was low stakes was shared by at least one center director, who explained, “This is where the industry is going, so I wanted to be involved at the point where it’s sort of consequence-free, to a certain extent. You can learn about it and there’s not anyone coming out and saying, “Oh, guess what? You’re a ‘2.’ We get an opportunity to learn and be involved.”

Second, some administrators thought it would be misleading to publish ratings when the Hybrid Rating Matrix (specifically, the ERS element) had changed over time and might continue to change. These leaders considered it important to accompany the publicizing of the ratings with a public information campaign that would explain what the ratings represent. But, as one administrator summed up the issue, the Consortium “would not want to turn around in a year and say, ‘Okay, now [a rating of four] means something different.’”

Third, a few administrators felt that the system, particularly the rating matrix, is “geared toward” publicly contracted/center-based programs because they are mandated by law or contract to comply with and/or conduct several of the components (such as maintenance of more protective staff-child ratios, use of DRDPs and the Ages and Stages Questionnaire [ASQs], and hiring higher qualified staff). Without a requirement to participate or financial incentives to attain these factors, it was doubtful that private centers serving fee-paying families would agree to participate. So, recruiting the broader spectrum of providers to participate “still remains a

challenge," said one administrator. "It's a challenge until every single licensed place in the state has to be part of a system that gets rated. It's a challenge, because people can say well, no rating is better than a bad rating." But, it should be noted that any rating above Tier 2 means that the site has exceeded minimum licensing requirements; this may not be clear to prospective participants.

Fourth, the frequency and approach to ratings could negatively impact sites that have been the most willing to be rated. For example, if a site gets a low rating, it will likely not have the chance to obtain a new rating for another two years. According to one leader, "So they're kind of stuck for two years, which seems unfair to them." In contrast, if sites receive provisional ratings (that is, temporary ratings assigned to sites that are waiting to receive ERS or CLASS observations and therefore do not have valid scores on those elements), they might have the opportunity to make improvements before their final rating. In one Consortium, for example, a site might receive a provisional rating because it is waiting on ERS or CLASS assessments. While it awaits the assessments, it might implement the ASQ or DRDP. When it is assessed at a later point, the entity in charge of the ratings asks the site personnel whether anything has changed since the provisional rating that they would like to have revisited (for example, that they have implemented the ASQ or DRDP). If anything has changed, the agency will go back to the site and revisit that piece before determining a final rating.

Finally, several QRIS administrators emphasized that parents in publicly contracted programs—who represent the bulk of the parents whose settings are participating in the pilot—have their child care choices narrowed to programs for which they are eligible. So parents who learn that their program's rating is not as high as they would like may not actually be able to act on this information and select a higher quality program for their child, given limited availability of contracted programs. "It's just so counterintuitive to what the intent of a QRIS really is," explained one administrator. Other leaders feared that, in the absence of more efforts to help families afford quality care, the publicizing of the ratings would not create consumer demand for a higher level of quality even among private fee-paying families.

Some QRIS administrators were enthusiastic about providing ratings to the sites because it afforded early learning and care providers an opportunity to improve and grow.

Several QRIS administrators underlined the potential benefits of publicizing ratings, stating that they believed that sharing ratings could reassure parents and offer providers an opportunity for learning and growth. As one leader stated:

> Providers can be "at the forefront of this statewide system and …they can use their individual ratings as a marketing tool. And talk to parents about what it is that they've been doing and why they were doing it. And the benefits that their families and the kids are getting from participating. And be able to have this… differentiator [from] a lot of other providers. And to say that they're committed to this quality process." Another administrator noted that publicizing the ratings meant that the "programs could really feel a sense of accomplishment. And they can see that [the QRIS administrative agency is] working to promote them."

And, as indicated above, despite the concerns about publicizing the ratings during the pilot phase, as of the end of 2015, according to the state Implementation Team, all of the pilot counties have in some way publicized the ratings. A majority of them are on websites; in some counties participating providers have a window or lawn sign featuring the program's rating. In other counties the publicized rating highlights particular details, such as the sites' strong points on particular portions of the RTT-ELC Hybrid Rating Matrix.

Because the Consortia had not yet implemented the publicizing of the ratings during the course of this evaluation, it was not yet possible to assess their impact on parental awareness and program improvement.

The Consortia reservations about publicizing ratings during a period when the rating matrix was still changing and the participation was voluntary are fully understandable. However, because the publicizing of the ratings took place largely outside the timeframe of this evaluation, we were not able to assess the degree to which ratings might impact both parental choice of quality programs and provider focus on program improvement. Now that the ratings are beginning to become available, it will be interesting to watch whether the ratings motivate providers to improve quality or help families access better care.

Sustainability

QRIS administrators also were asked whether their Consortia would be able to sustain all aspects or elements of their RTT-ELC-funded QRIS after RTT-ELC funding ends.

Some Consortia feel confident about sustaining their QI activities through collaborations and access to new resources after RTT-ELC funding ends. For example, in one Consortium, a longstanding history of QI initiatives and programs that had a structure, scope, and mission similar to (or more extensive than) that of the RTT-ELC QRIS, coupled with substantial local funding, meant that the end of RTT-ELC funding was not expected to impact sustainability at all. When a second Consortium first began implementation of the RTT-ELC QRIS, an administrator thought that "they did a really great job looking at where the cost burdens would be when we started this [implementation of RTT-ELC QRIS]. Among those 'cost burdens' was the assessment cost. And if you can systemically figure out a model that minimizes that output, then you have a cost effective model and I feel like we were able to do that."

In many Consortia, several QI activities will be sustained because of other funding streams and initiatives. For example, one leader said that her Consortium has always provided a great deal of on-site support, training opportunities and training materials, and resources for programs on the DRDP. Another Consortium would be able to sustain its ASQ work because it is located in a Help Me Grow county. In at least one Consortium, strong partnerships with local colleges would allow for individual professional development opportunities to be sustainable.

One method of ensuring sustainability is to increase capacity among various programs in the community. For example, in one Consortium, administrators had been trying to build capacity by helping to develop coaches who are located within—and supported by—larger agencies and by increasing the capacity of other community agencies, such as Head Start, to provide in-house training.

Offering a different take on sustainability, another Consortium leader noted that its administrators had designed its system with sustainability in mind from the onset. For example, they did not offer stipends for professional development because they wanted to establish a culture in which early learning staff were working to improve the level of quality in order to make a difference in the lives of the children they served, as opposed to doing so to receive a monetary stipend.

The CSPP QRIS Block Grant and the I/T QRIS Block Grant are expected to have a significant impact on sustainability, helping to replace the federal RTT-ELC grant funds.

Since 2012, the 17 Consortia in 16 counties, as shown in Exhibit 2.4 below, have received a total of more than $57 million in federal grant awards for program quality assessment, rating, monitoring, and program quality improvement activities.

Exhibit 2.4. RTT-ELC Grant Awards to 17 Consortia

County	Administrative Agency	Grant Award
Alameda	First 5 Alameda	$3,720,756.00
Contra Cosa	First 5 Contra Costa	$2,062,928.00
El Dorado	First 5 El Dorado	$689,800.00
Fresno	Fresno County Office of Education	$2,989,080.00
Los Angeles	County of Los Angeles Office of Child Care	$6,699,113.62
Los Angeles	Los Angeles Universal Preschool	$7,888,024.52
Merced	Merced County Office of Education	$1,242,260.00
Orange	Orange County Office of Education	$6,695,017.48
Sacramento	Sacramento County Office of Education	$3,822,234.48
San Diego	First 5 San Diego	$6,761,628.88
San Francisco	First 5 San Francisco	$2,020,560.00
San Joaquin	First 5 San Joaquin	$2,626,377.52
Santa Barbara	First 5 Santa Barbara	$1,566,858.00
Santa Clara	First 5 Santa Clara	$4,130,736.00
Santa Cruz	First 5 Santa Cruz	$1,187,145.48
Ventura	First 5 Ventura	$2,612,950.00
Yolo	First 5 Yolo	$850,800.00
Total Grant Awards to 17 Consortia		**$57,566,269.98**

Looking ahead as these funds are about to expire, several QRIS administrators expect that CDE's new CSPP QRIS Block Grant will definitely help sustain the QRIS. According to CDE's website, "Senate Bill 858 (Chapter 32, Statutes of 2014) authorized $50 million of Proposition 98 funds for Quality Rating and Improvement System (QRIS) Block Grant for the support of local early learning QRIS that increase the number of low-income children in high-quality state preschool programs that prepare those children for success in school and life." As shown in

exhibit 2.5, all 17 Consortia in the RTT-ELC pilot have received funds from Phase I of the QRIS Block Grant, receiving a total of $39 million in 2015-16.

Exhibit 2.5. Fiscal Year 2015–16 California State Preschool Program (CSPP) QRIS Block Grant RFA Funding Results, Priority 1A

County	Administrative Agency	Amount
Alameda	Alameda County Office of Education	$2,310,787
Contra Costa	Contra Costa County Office of Education	$1,307,937
El Dorado	El Dorado County Office of Education	$253,960
Fresno	Fresno County Office of Education	$2,569,776
Los Angeles	Los Angeles County Office of Education	$15,479,004
Merced, Madera, Mariposa	Merced County Office of Education	$1,061,940
Orange	Orange County Department of Education	$3,321,599
Sacramento	Sacramento County Office of Education	$2,177,102
San Diego	San Diego County Office of Education	$4,321,096
San Francisco	San Francisco Unified School District	$1,168,385
San Joaquin	San Joaquin County Office of Education	$1,160,423
Santa Barbara	Santa Barbara County Office of Education	$658,788
Santa Clara	Santa Clara County Office of Education	$1,729,947
Santa Cruz	Santa Cruz County Office of Education	$382,198
Ventura	Ventura County Office of Education	$955,075
Yolo	Yolo County Office of Education	$349,510
Total for 17 Consortia in QRIS Pilot		**$39,207,527**

SOURCE: California Department of Education: http://www.cde.ca.gov/fg/fo/r2/csppqris1516result.asp

According to several administrators, the CSPP QRIS Block Grant will help them sustain some level of engagement, albeit not always the same degree of engagement as with RTT-ELC. One issue, for example, is that only 20 percent of the CSPP QRIS Block Grant can be spent on program assessment, rating, and monitoring; by contrast, a significant proportion of the $57 million in RTT-ELC grant money has been used for these activities to date. As one administrator noted, work would continue after the federal funding ended; however, “at what service model we can hold people to, and what we can afford in terms of assessment, remains to be seen.”

This same administrator explained that “we’re pretty comfortable that by merging CSPP QRIS Block Grant dollars and AB212 dollars, we will continue to have a [QRIS] presence for our CSPP programs.” Another leader noted, “So we know that for our CSPPs, [sustainability] is a no brainer [because of the CSPP QRIS Block Grant]. They will continue QRIS as it is.” In another Consortium, CSPP funding will be used to fund additional early learning mentors. Because Head Start had done the same, the Consortium may actually expand the number of mentors in its area after RTT-ELC funding ended.

In another Consortium, leaders projected that coaches who were "heavily subsidized" by an administrative agency as of 2015 could be supported (at least in part) by other funding streams, such as the CSPP QRIS Block Grant, should the sites continue using the same approach. In another Consortium, administrators noted that there would be new ERS and CLASS assessors as part of the CSPP QRIS Block Grant.

The First 5 California IMPACT grant is also expected to help promote QRIS sustainability.

QRIS administrators also expected the First 5 California IMPACT grant to help support sustainability. The RFA for the First 5 California IMPACT grant was issued in July 2015 and the final funding list for Phase I was posted in November 2015.[5] According to the IMPACT RFA, the purpose of the grant is to develop a network of QI systems in order to support more sites in achieving high standards of quality and also to provide information to parents to enable them to better promote their children's development and learning. First 5 California has pledged $190 million for the First 5 IMPACT grants over five years (i.e., until 2020).

A key goal of the IMPACT grant is to offer more QI support to the private programs, which have to date had less reason to participate in the QRIS and to align with other QI funding streams to help achieve that goal.

First 5 county commissions have priority for the IMPACT grant and partners must include, at a minimum, the CSPP QRIS Block Grant lead, local education agency(ies) (such as the County Office of Education or school district), the RTT-ELC lead, Local Child Care and Development Planning Councils, local R&Rs, institutions of higher education, and other local entities as determined by the lead agency.

As shown in exhibit 2.6, all 17 Consortia in the RTT-ELC pilot received base-layer funding from Phase I of the First 5 IMPACT RFA. And, as indicated above, as of spring 2016, all 58 counties now participate in the First 5 IMPACT Grants.

The following settings and providers are eligible for IMPACT grants: (1) licensed centers; (2) licensed family child care providers; (3) license-exempt providers; and (4) alternative settings (that is, community and home-based settings that are not child care or preschool sites and provide early learning and school-readiness services to parents and young children).[6] Funding cannot be used for direct services; however, it can be used to support improving the quality of direct services (for example, coaching, resources or materials, stipends, provider/teacher technical assistance).

Exhibit 2.6. First 5 IMPACT RFA, Phase 1 Base-Layer Funding, Fiscal Year 2015–16 to 2019–20

County	Administrative Agency	Amount
Alameda	First 5 Alameda County	$1,387,186
Contra Costa	First 5 Contra Costa Children and Families Commission	$939,505
El Dorado	First 5 El Dorado	$204,312
Fresno	First 5 Fresno County	$1,452,911
Los Angeles	First 5 Los Angeles	$5,000,000
Merced (and Mariposa)	First 5 Merced County	$579,969
Orange	Children and Families Commission of Orange County	$1,936,553
Sacramento	Sacramento County Office of Education	$1,699,976
San Diego	First 5 Commission of San Diego	$2,937,013
San Francisco	First 5 San Francisco	$565,418
San Joaquin	First 5 San Joaquin	$841,471
Santa Barbara	First 5 Santa Barbara	$513,900
Santa Clara	First 5 Santa Clara	$1,468,902
Santa Cruz	First 5 Santa Cruz County	$345,461
Ventura	First 5 Ventura	$793,033
Yolo	City of West Sacramento	$289,935
Total for 17 Consortia in QRIS Pilot		**$20,955,545**

SOURCE: First 5 California

Thus, the First 5 IMPACT grant might help accomplish the RTT-ELC goal to include various types of settings and programs. According to one QRIS administrator, "My hope is that with First 5 California IMPACT [funds], we should be able to engage more non-CSPP providers that otherwise would have been excluded from participating in any kind of quality improvement activities."

Despite potential new funding streams available to the state, many Consortia think they will have to scale back on some of the activities and/or scope of enrollment in some child care settings after the end of the RTT-ELC grant funding.

The CSPP QRIS and I/T QRIS Block Grants, assuming they continue, and the First 5 IMPACT Grant, at least for five years, will ensure significant continued funding for QRIS activities, especially in California State Preschool sites. However, the majority of Consortia administrators noted that would either have to scale back on some of their activities and/or limit enrollment in particular settings after the RTT-ELC funding ends.

As one administrator said, once RTT-ELC funding ended, the Consortium would have minimal funds to conduct ratings in non-State Preschool settings, and it would have *no* funding for QI other than the CSPP QRIS Block Grant funding, and most of those monies would go directly to

the agencies operating State Preschool. After the RTT-ELC funding ended, this Consortium was going to reduce costs by offering coaching on a fee-for-service basis as well as an online coaching model.

Finally, it will be important to watch whether or how much the structuring of the CSPP QRIS Block Grant and the awards available to each Consortium change over time. Now that all 58 counties are participating, unless the state increases the total funding for the block grant, each Consortium's allocations (and therefore its QI grants and quality rewards) may decrease as more counties beyond the original 16 participate. That is, as more State Preschool programs get rated, the money for the CSPP QRIS Block Grant will be spread out across those additional programs.

In addition, in 2015, CSPP programs could receive both IMPACT's Quality Rewards and Quality Improvement support; however, after 2015, the approach will change. Programs at Tiers 4 and 5 will be eligible to receive IMPACT's Quality Rewards and programs at Tiers 1, 2, and 3 will be eligible to receive IMPACT's Quality Improvement support.

When asked about how other potential funding streams beyond the RTT-ELC, such as the First 5 IMPACT grant and private philanthropy, would foster sustainability, several administrators noted that although this support would help, it would likely not be sufficient to sustain all of the components implemented as part of RTT-ELC, and Consortia would have to scale back. For example, one Consortium leader shared that there are "pockets where philanthropy is helping. But, that doesn't help us go to scale. I think the truth is that if scaling is a goal, we have to reduce our intensity. It's just one of our challenges here, is really coming to terms with how to make a good decision about depth versus reach."

One way to scale back is to identify those QI activities that are most critical and look for ways to reduce costs and/or discontinue other QI activities. For example, according to one leader, her Consortium has prioritized different QI activities (coaching, educational incentives, and outreach) and "rearranged funds in a way" that they could be sustained. This leader added, "And that's what I think the biggest lesson has been from this experience—[it] is being able to identify what you need, articulate it, and then put it into the back end of the program." Another leader said that although "the best way to ensure that things are well implemented is through a coaching model," a coaching model is very time-consuming and costly; therefore, the Consortium might consider different approaches to coaching (such as virtual coaching) in the future.

Some Consortia plan to scale back by focusing on serving and enrolling a subset of providers or limiting some QI activities to particular sites or cohorts. For example, one Consortium will provide coaching to sites in Tiers 1, 2, and 3 (and not to sites in Tiers 4 and 5). Another Consortium will "not water down the entire training and TA effort," but will just offer to fewer sites. In another Consortium, the first cohort was scheduled to stop receiving QI supports at the end of June 2015 and both cohorts would stop receiving QI supports after the end of RTT-ELC funding. In a fourth Consortium, the State Preschool sites will be sustained by the CSPP QRIS Block Grant; administrators were less clear about the future of sites that are only partially funded with CSPP QRIS Block Grant funds and sites that receive no such funds. This same leader hoped that the First 5 IMPACT funds could help sustain some of this latter group.

Another way to support sustainability is to revisit and modify the Hybrid Rating Matrix, particularly concerning the ERS element. Some administrators felt that the changes made in mid-May 2015 would be sufficient, while others mentioned additional potential changes in the future. For example, according to one Consortium leader, "If we continue to have this high stakes assessment tool, we won't be able to sustain [the work]." The same administrator voiced concerns that there is a new ERS assessment tool and it would not make any sense to continue with the existing one. Yet another Consortium leader said, "Like a lot of other counties, there's a lot of interest in either removing the ERS scale, because it's not financially viable or somehow reducing in the requirements or limiting it in a way that would make it financially viable…. It's just so hard to maintain reliable assessors and to complete those assessments in the manner that it's written currently."

In at least one Consortium, the entire RTT-ELC-funded program will end after RTT-ELC funding ends. One administrator shared that the CSPP participants will transition over to the CSPP QRIS Block Grants, but that services to other types of providers will end after June 2016.

Fee-for-services could foster sustainability.

A couple of Consortia are considering fee-for-services as one way of fostering sustainability. For example, as noted above, after the RTT-ELC funding ends, one Consortium is planning to offer coaching on a fee-for-service basis. Another Consortium is considering a fee-for-services approach for ratings for privately funded programs that wanted to have external ratings because "there would be providers there who would be willing to pay to get rated just so they can say, I'm in a QRIS and I'm a Tier 5.'" But given the current and impending demands of other funding sources, this Consortium leader was not certain the administrative agency had sufficient human capacity to add more responsibilities to its plate.

A robust QRIS requires significant investments, including sufficient administrative funding.

Many QRIS administrators argued that a robust QRIS that enhances quality is only possible if significant investments are made. An administrator in one Consortium noted that given the proportion of RTT-ELC funding to local funding, sustainability would not be an issue; however, this same administrator noted that quality is an expensive undertaking. In another Consortium, an administrator said that the cohort model was one of the best ways to get child care professionals to attain college degrees, but added, "They're really expensive. If we are committed to degree attainment, I don't think we're going to be able to get there on the cheap." As another administrator shared, "There needs to be some shift in where all the money goes, and we need more money in coaching, and professional development, and QI and things like ERS. The current cost of that makes that pretty impossible, I think, to do well, unless you're in a county that has some agency or some county coffers that are going to fund it… [you cannot] start shifting the burden on Prop 10 First 5 to cover these costs when we [First 5s] were all facing significant decreases in our funding." An administrator from another Consortium echoed this comment, noting, "A lot has been put on the shoulders of First 5 in each county, without giving consideration to the shrinking investment, or shrinking dollars from Prop 10 and [the fact that] local contributions are not necessarily directed toward this [QRIS] effort."

When one Consortium conducted a cost analysis of their model and approach, “it was a pretty eye-opening experience for all of our team--even those who are very committed to keeping the model the same.” This Consortium was using the findings from that cost analysis to brainstorm about how it could provide coaching and technical assistance in different ways that are more cost effective and yet still maintain the quality of its model.

A couple of administrators noted the lack of permanent funding from the state to support the QRIS effort. As one administrator shared, “When you talk [about] sustainability, I have a hard time answering that, because for everything we’re going from grant to grant, from program to program, to initiative to initiative. Can we sustain it? Yes, because we see what else is coming and we can always look at that to sustain. But is it a permanent source of funding that’s going to be here to stay to support this effort forever? We don’t have one [permanent source of funding for QRIS] as a state.” Finally, two leaders made note of the administrative costs necessary to run a QRIS.

Administrators shared concerns about how some new initiatives such as the CSPP QRIS Block Grant might support a rating system but not sufficient funds for QI supports. As one leader noted, “It doesn’t make any sense for the state to impose a rating system on those programs without providing any additional resources to help those programs get up to a higher tier level, in terms of their measured quality.” This leader argued that many Consortia do not have local resources devoted to QI, and yet “they are being asked to assess and rate programs and improve quality without any resources to improve quality. So I think the state really needs to not just be in the rating business, but also step up in the quality improvement area, which is where most of the dollars are needed.”

Changes to the RTT-ELC matrix could potentially influence the CSPP and I/T QRIS Block Grants and the First 5 California IMPACT grant.

One of the primary purposes of the RTT-ELC pilot QRIS was to validate the Hybrid Rating Matrix and rating system. So, as in many states with new QRISs, it is not surprising that some changes in the rating matrix are still under consideration. As one leader shared, “And I don’t think any of us think the matrix is done. You know, this isn’t the end-all, be-all, written-in-stone type of quality matrix, because there are things that we want to add to it, like … whole family engagement, in particular.”

Several administrators noted how the aforementioned changes to three-point level and five-point level of the ERS element would affect or align with the recently implemented CSPP QRIS Block Grant or the First 5 California IMPACT grant. As one administrator noted, “When we’re looking at the CSPP sites that are coming on board as [CSPP QRIS] Block Grant sites, it makes sense to us that we want that baseline rating with the ERS [external] assessment, because it’s so high stakes now.” Another administrator noted that there seems to be a push from the new I/T QRIS Block Grant and First 5 California’s IMPACT grant “to really start to codify the matrix, as it exists,” when throughout the process, many of the Consortium leaders had been waiting to collect enough data to inform discussions about which aspects of the rating matrix might need to be modified or adjusted.

Finally, another Consortium leader shared, “The task upon us is to decide what the matrix will look like after RTT-ELC sunsets in 2016. Will we have the same matrix?...Whatever changes we make in that matrix have a direct impact on the CSPP QRIS Block Grant and First 5 IMPACT.” So the question is “what kind of matrix and quality continuum framework are we able to maintain”.

Summary

California is near the completion of its RTT-ELC QRIS pilot and, as would be expected of a pilot, the system continues to grow and change. For example, the Consortia have done a considerable amount of work implementing the Hybrid Rating Matrix and fine-tuning and modifying it based on lessons learned.

The majority of the Consortia had implemented most of their planned QRIS activities as of summer 2015. Site and classroom assessments were one integral piece of these planned activities. The majority of the Consortia were following the RTT-ELC’s requirements for the frequency and sampling of classroom assessments, although several went beyond the minimum requirements and were conducting annual observations or expanding the sample size of observed classrooms at a site. Challenges related to finding, training, and retaining qualified classroom assessors, particularly for the ERS lessened between 2014 and 2015 because the Consortia that had initially struggled had found ways to make observations more affordable and manageable.

In addition, most Consortia had reached their target goals for the total anticipated number of QRIS participants as of summer 2015. Although the pilot system did include privately funded centers and FCCHs, most of the participants were publicly contracted programs such as State Preschool and Head Start, and, as anticipated, included only a small fraction of the licensed centers and an even smaller percentage of licensed FCCHs in the counties participating in the RTT-ELC QRIS. By comparison, according to the online QRIS Compendium, updated through 2015, in states such as Illinois, New Hampshire, and Oklahoma, 100 percent of licensed centers and FCCHs participate in the QRIS. In effect, obtaining a license automatically awards a program one “star” in these states. In Tennessee, where 91 percent of centers and 88 percent of FCCHs participate in the QRIS, all licensed programs through its Department of Human Services are required to be rated, and those that receive at least one star may choose to participate in the Star Quality Program and receive QI supports and other incentives. In Wisconsin, participation is mandatory for programs receiving federal and state subsidies, with 82 percent of licensed centers and 75 percent of licensed FCCHs participating.

During this QRIS pilot, the Consortia have largely focused on conducting quality assessments and QI activities to help improve program quality. Although ratings have been provided internally to providers, most Consortia had not publicized ratings by summer 2015, and hence we were not able to evaluate the impact of publicized ratings on quality improvement and parent access to quality programs. Many of the Consortia were still working on a plan for rolling out the ratings, though, according to the state Implementation Team, all publicized ratings by December 2015.

In terms of validation, it is important to reiterate that California’s QRIS is not one uniform system. For example, although all of the Consortia use a common, five-tiered Hybrid Rating

Matrix, each Consortium can elect to block Tier 2 and make local modifications to elements within Tier 5. This allowance for local modifications, in addition to the changes at three-point level and five-point level of the ERS element of the rating matrix in May 2015, shed light on elements that particular Consortia deem most important but they also complicate attempts to validate the system.

Finally, it may be difficult to expand—and possibly even to sustain—the work done as part of the QRIS pilot without changes in the requirements for participation in the system and/or other financial resources to help rated sites improve. Although the CSPP QRIS Block Grant and First 5 IMPACT grants are expected to make large contributions to ECE in California, and the 17 pilot Consortia in particular (a total of more than $60 million to the 17 pilot Consortia from both IMPACT grants and CSPP QRIS Block Grants), it is unclear whether more private providers and voucher provides will participate in the QRIS system without a requirement that they do so. In addition, now that the pilot counties appear to be publicizing ratings, it will be important to assess the extent to which ratings either inform parents or motivate quality improvement. In a voluntary system, the issue of participation might be particularly tenuous for private providers if they are expected to be assessed without receiving sufficient QI supports.

In this chapter, we have emphasized the limited nature of the types of providers participating in the 17 RTT-ELC pilot QRISs. In a state as large and diverse as California, wisdom suggests the benefits of conducting a pilot before considering statewide implementation. However, by not having the authority to require all licensed programs to participate in the pilot or the resources to encourage their participation within the 16 counties, the pilot QRISs have focused on publicly contracted programs with limited outreach to private centers or FCCHs with fee-paying parents. As a result, the evaluation results may not tell us much about how a QRIS with a broader spectrum of programs participating would function. If more private programs participated, it might be much more difficult and expensive to conduct the kind of program quality assessments that have been implemented by the Consortia. But if all licensed programs received a publicized one star as a condition of licensure, there might be far more incentive for private programs to participate in the system to obtain whatever QI supports were available to help them rise to a higher publicized level.

Although there are issues to be addressed, as noted above, the Consortia, within the limitations of the RTT-ELC grant, have accomplished a great deal in the four years since the grant was awarded and the five years since the CAELQIS Advisory Committee made its recommendations. Many of the QRISs in other states have taken more than a decade to reach full implementation, and even some of the best-known systems, such as North Carolina's, regard refinement of the system as an ongoing effort. The fact that all 58 counties are now participating in some aspects of the QRIS provides tangible evidence of the growing interest in the system.

Chapter 3. Validity of the QRIS Ratings

Key Findings

This chapter explores the validity of the California QRIS ratings in alignment with the intended purposes of the ratings. We examine the measurement properties of the ratings, and assess the validity of the ratings as a measure of program quality and for the purpose of identifying programs that support child learning and development outcomes.

- Analyses of rating elements and their relationship to overall program ratings indicate that QRIS ratings do not represent a single dimension of quality, suggesting that the best way to disseminate the QRIS ratings may involve presenting more detailed scores rather than a single, overall rating.
- The analyses find that the California QRIS ratings are positively related to measures of instructional support and the quality of adult-child interactions, but are not associated with other aspects of quality such as daily routines or curriculum planning and assessment.
- 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, we found only a small positive relationship between tier ratings and one of four outcomes measured. These mixed results are consistent with other states' QRIS evaluations.
- Analyses suggest that validity results may be slightly better for QRIS ratings calculated by averaging the element scores, rather than the current approach used in California of summing the total points accumulated across elements.

The results of the validity analyses should be considered in light of the limitations of the study sample, which had limited variability in QRIS ratings and element scores and is dominated by programs with standards-based public funding. Findings could differ with a more diverse sample of programs.

This chapter explores the validity of the California QRIS ratings in alignment with the intended purposes of the ratings. We highlight results from the half-term report regarding the measurement properties of the ratings and the validity of the ratings as a measure of program quality, and present new results regarding relationships between ratings and child outcomes. The chapter addresses the following research questions:

- RQ4. How effective are the California Common Tiers' structure and components/elements at defining and measuring quality in early learning settings?

- RQ5. To what extent do the graduated elements and tiers correspond to graduated increases in child outcomes, including (but not limited to) children's learning, healthy development, social-emotional health, and school readiness?
- RQ 6. To what extent can the Consortia's local QRIS be streamlined and still result in the same program quality level and child outcomes? What common elements of the Hybrid Rating Matrix and Pathways are most important to include?

We begin the chapter by examining the distribution of QRIS ratings and element scores as well as the measurement properties of the QRIS ratings among programs with full ratings[7] at the start of the study. To determine how well the ratings function as a measure of program quality, we first examine the distribution of QRIS ratings and element scores as well as the program characteristics that predict QRIS ratings. Next, we examine the internal consistency of the ratings—the extent to which the individual rating elements that comprise the rating are related to each other and to the overall rating. These results also were presented in the half-term report.

After exploring the measurement properties of the ratings, we assess the validity of the ratings for the purpose of differentiating programs according to the quality of their services by comparing programs' QRIS ratings with independent measures of quality to see how closely they are aligned. One way to demonstrate this type of validity is to observe significant increases in program scores on valid, research-based measures of early childhood program quality with each QRIS rating increase. These findings also were discussed in detail in the half-term report.

Next, we present a new set of analyses examining the validity of the QRIS ratings for the purpose of identifying programs that support child learning and development outcomes. Specifically, we assess whether centers' QRIS ratings or element scores are associated with children's scores on assessments of early literacy, mathematics, and executive function skills in the spring of the program year, controlling for their fall scores. If children's scores on developmental assessments increase significantly with each QRIS rating increase, it would suggest that the ratings can be used reliably for this purpose.

Finally, we explore alternative approaches to calculating the ratings using California's existing element scores, to determine if any alternative approaches generate ratings that have more optimal results in our validation study analyses. Specifically, we examine the relationship between the alternative rating approaches and program quality measures (as presented in the half-term report) as well as child outcomes. We then compare the results to California's QRIS rating approach.

This chapter examines the measurement properties of centers and family child care homes (FCCHs) separately, and presents classroom quality and child outcome results for center-based programs only. The QRIS ratings use different criteria for centers and FCCHs, and the

---

[7] Programs with full ratings are defined as those with complete, nonprovisional QRIS rating data as of January 2014, when study sampling occurred. Some Consortia had assigned provisional QRIS ratings to programs that had begun the QRIS rating process but were on a waiting list to have an independent observer complete the QRIS rating data collection. For the study, each Consortium identified the programs within its local QRIS that had provisional ratings as of January 2014; these programs were not included in the analyses for this report.

distributions of ratings are very different for each program type, so the results of validity analyses combining both program types are difficult to interpret. However, results of models that combine centers and FCCHs are presented in appendix 3 for the reader's reference. The chapter does not present results of these analyses for FCCHs because there are too few in the analysis sample to support reliable analyses with FCCHs only.

As described in the analysis approaches box on the following page, the programs with full QRIS ratings that are included in the study analyses tend to be high in quality and differ from the programs without full ratings that were not included in the study (see appendix 1A). Therefore, results of the analyses in this chapter may not apply to the broader range of programs participating in the QRIS. In addition, the analyses for child outcomes include controls for some program and child background characteristics, but any observed differences between children in programs at each rating level could potentially be due to factors not accounted for in the study, such as family income and parent education. Thus, results must be interpreted with caution.

How Well Does the QRIS Perform as a Measure of Quality?

To determine how well California's QRIS ratings function as a measure of program quality, we analyzed QRIS rating data, including rating levels and element scores, from programs across the state with full QRIS ratings as of January 2014. The analyses included an examination of the distribution of ratings and element scores, the characteristics of programs that predict QRIS ratings, the internal consistency of the ratings, and how element scores relate to each other and to the overall rating.

Among programs with full ratings, the distribution of QRIS ratings is limited and does not span all five possible QRIS rating levels.

Results indicate that, first, the distribution of ratings in the limited sample of fully rated programs is limited; it does not span all five possible QRIS rating levels. Among the sample of 472 programs with full ratings, no programs were rated at Tier 1 using California's QRIS criteria. In contrast, 19 percent of programs with provisional ratings were rated at Tier 1, suggesting that programs do not complete a full rating until they are able to earn enough points to achieve Tier 2. Ratings of fully rated programs were generally high, with half of all sites rated at Tier 4 or higher. This may be due to the population of programs participating in the system: as a voluntary system, programs that might score lower have little motivation to become involved. In fact, many of the fully rated programs are State Preschool or Child Signature Program (CSP) sites—many of the fully rated programs are State Preschool or Child Signature Program (CSP) sites—programs with specific quality requirements—and many have been participating in quality improvement efforts for many years (prior to Race to the Top-Early Learning Challenge [RTT-ELC] funding and the development of the Hybrid Rating Matrix) and thus have had the benefit of significant professional development and quality improvement resources. However, relatively few programs (8 percent) were rated at Tier 5, indicating that the Tier 5 criteria set a high bar.

Analysis Approaches

Below we describe the methods used to address four main questions, drawing from different combinations of data sources.

How Well Does the QRIS Perform as a Measure of Quality?

- Rating Distributions: Description of the number of programs at each QRIS rating level and element score level
- Predictors of Ratings: Ordinal logistic regression analysis indicating if program characteristics predict ratings
- Internal Consistency: Cronbach's alpha statistics assessing the extent to which the QRIS rating measures a single latent construct of program quality
- Element Scores and Ratings: Descriptive and correlational analysis describing how the element scores relate to each other and the overall rating

Data and Sample: QRIS rating data, including ratings and element scores, from 472 programs across the state with full QRIS ratings as of January 2014.

How Well Does the QRIS Differentiate Programs Based on Observed Quality?

- QRIS Ratings and Observed Quality: Analysis of variance (ANOVA) models, examining the average scores on independent measures of observed quality by QRIS rating level or element score level for centers only. ANOVA models indicate whether the average scores differ significantly by rating or element score level. ANOVA models were not conducted with FCCHs due to the small sample, although differences in scores are examined descriptively by QRIS rating level for FCCHs.

Data and Sample: QRIS rating data and classroom observation data with the Classroom Assessment Scoring System (CLASS) and the Program Quality Assessment (PQA). The analytic sample includes 139 centers and 20 FCCHs with CLASS scores, and 140 centers and 27 FCCHs with PQA scores.

How Well Does the QRIS Predict Children's Outcomes?

- QRIS Ratings and Child Outcomes: Multilevel regression models tested associations between QRIS ratings or element scores and children's developmental outcomes. The models test whether children enrolled in higher rated programs perform better on assessments of language and literacy, early mathematics, and executive function skills in the spring of the program year than the fall. The regression analyses use multilevel modeling to account for the grouping of children within sites, and the models control for child developmental skills in the fall and a variety of site and child characteristics.

Data and Sample: QRIS rating data and direct child assessments, including measures of preliteracy skills (Woodcock-Johnson Letter-Word Identification subtest and Story and Print Concepts), mathematics skills (Woodcock-Johnson Applied Problems subtest), and executive function (Peg Tapping task). The analytic sample included 113 center-based programs and 1,501 to 1,611 children, depending on the outcome studied.

How Do Alternative Rating Approaches Affect the Distribution and Validity of Ratings?

- Calculation of Alternative Ratings: QRIS ratings are calculated using a variety of calculation approaches, using the element score data collected for QRIS ratings, and the resulting ratings are compared across approaches.
- Alternative Ratings and Observed Quality: ANOVA models assess whether average scores on independent measures of observed quality in centers differ by each alternative QRIS rating approach. The results of the ANOVA models are compared to determine which rating approaches best differentiate observed program quality.
- Alternative Ratings and Child Outcomes: Multilevel regression models tested associations between each alternative QRIS rating approach and children's developmental outcomes. The results of the multilevel regression models are compared to determine which rating approaches best differentiate child outcomes.

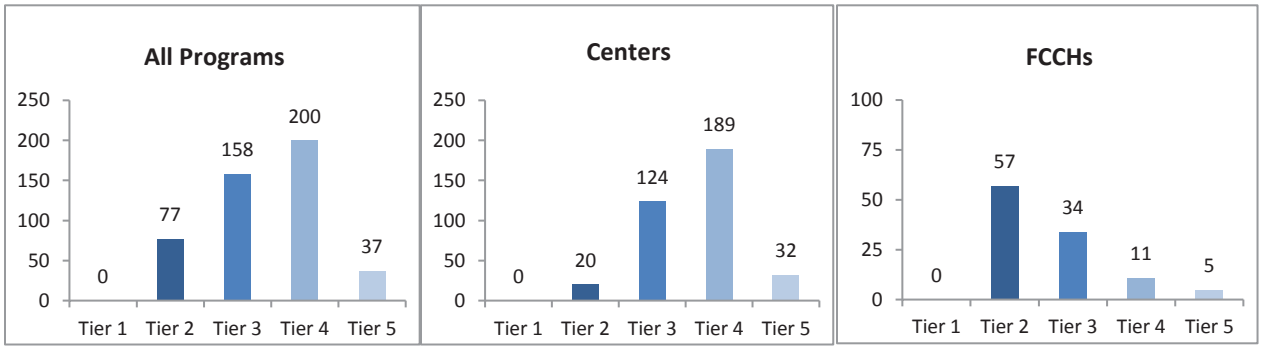
Data and Sample: Calculation of alternative ratings uses QRIS rating data from 472 programs. Observed quality and child outcome analyses with the alternative ratings use the same data and samples as the main QRIS rating analyses.

**Programs with full QRIS ratings in 2013 tend to be high in quality and differ from other programs participating in the QRIS in 2013 without full ratings (see appendix 1A). Therefore, results of the analyses in this chapter may not apply to the broader range of programs participating in the QRIS.

The distribution of ratings and element scores differs markedly for centers and FCCHs.

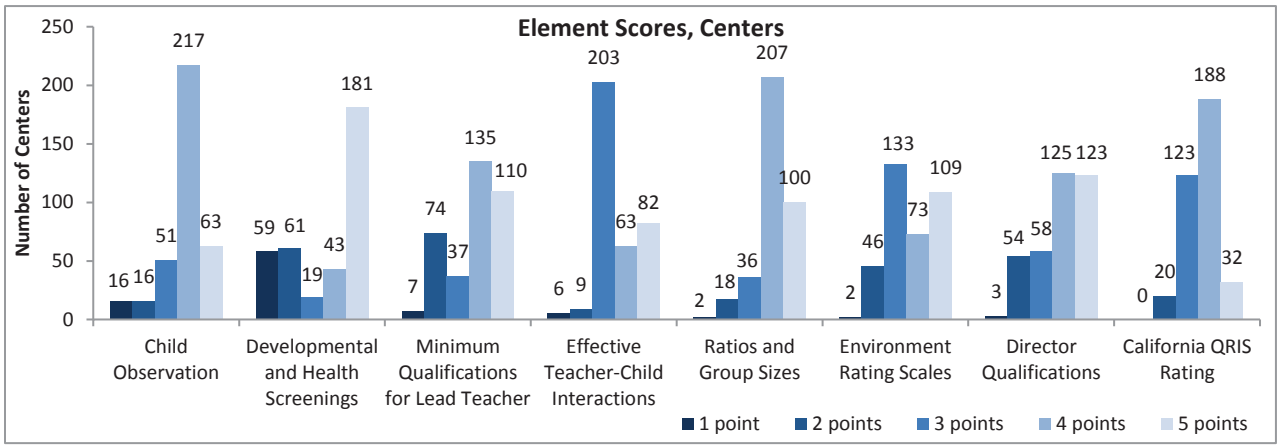
The distribution of ratings differs markedly for fully rated centers and FCCHs (see exhibit 3.2). Although the most common rating for centers is Tier 4, and 86 percent of centers were rated at Tier 3 or 4, the most common rating for FCCHs is Tier 2, and 85 percent of FCCHs were rated at Tier 2 or 3. Differences in ratings between centers and FCCHs may be partially explained by differences in the percentage of centers (96 percent), and FCCHs (43 percent) that are required to meet high-quality standards for State Preschool, CSP, or Head Start funding. It is not known if these patterns will be similar when a larger and more diverse sample of centers and FCCHs participates in the QRIS.

Exhibit 3.2. Distribution of California QRIS Ratings for All Programs, for Centers, and for FCCHs with Full Ratings in January 2014



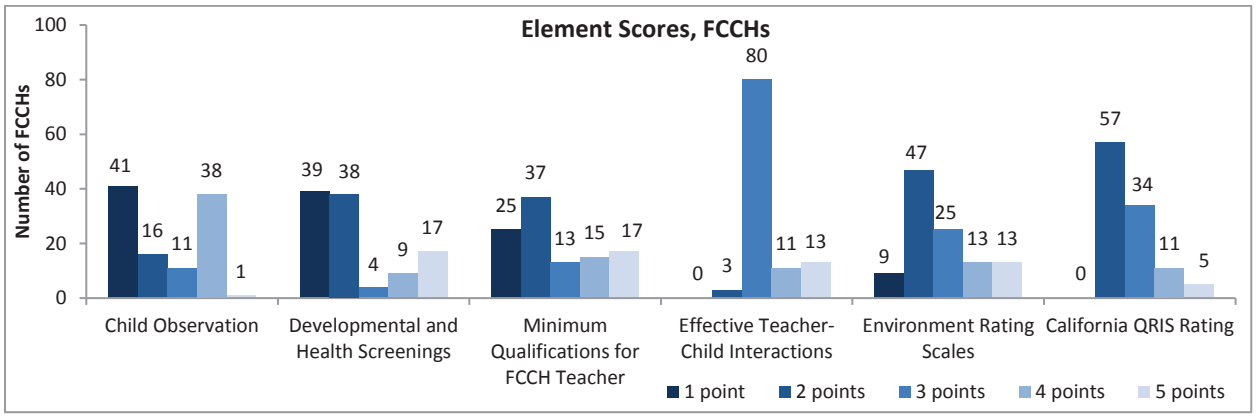
The distribution of element scores differs considerably between centers and FCCHs with full ratings, as shown in exhibits 3.3 and 3.4. Centers and homes differ in the distribution of element scores related to structural quality, which refers to easily measurable program characteristics that contribute to high quality, such as staff qualifications, curricula and assessment tools used by the program, adult-child ratios, and group sizes. Among centers, element scores tend to be high (more than 60 percent of centers received an element score of 4 or 5) for all of the element scores related to structural quality, including Child Observation, Developmental and Health Screenings, Minimum Qualifications for Lead Teacher, Ratios and Group Sizes, and Director Qualifications. This suggests that the rating criteria for these structural quality elements may not differentiate centers well, at least among the study sample of 363 fully rated centers that are not infant-only. Among FCCHs, element scores are more dispersed on the structural quality indicators (among those that are applicable to California QRIS ratings for FCCHs: Child Observation, Developmental and Health Screenings, and Minimum Qualifications for FCCH Provider) in comparison with centers.

Exhibit 3.3. Distribution of Element Scores: Centers



N = 363 centers, excluding 2 centers serving only infants

Exhibit 3.4. Distribution of Element Scores: FCCHs



N = 107 FCCHs

Centers and homes also differ in the distribution of element scores related to process quality, which refers to interactions between adults and children in classrooms, and includes constructs such as teacher sensitivity and instructional quality. Among centers, there is somewhat more variability among ratings in the elements that measure process, including effective teacher-child interactions (based on the CLASS instrument) and program environment rating scales (based on the Environment Rating Scale [ERS] instrument). Among the process quality elements, FCCHs have little variability on the effective teacher-child interactions element, but they do vary in scores on the program environment rating scales element score.

Among centers, CSP funding and Title 5 (State Preschool, General Child Care, or Cal-SAFE) funding are statistically significant predictors of California QRIS rating level.

CSP funding and Title 5 (State Preschool, General Child Care, or Cal-SAFE) funding are significant positive predictors of California QRIS rating level among centers, after controlling for Consortia, enrollment, programs serving infants and toddlers, and other funding streams. This is not surprising because requirements for public funding streams are closely aligned with

requirements for high scores on some QRIS elements. Consortium membership also predicted QRIS ratings, but none of the other program characteristics were significantly related to ratings for centers. However, only a limited number of predictor characteristics were included in the models due to data limitations, so these results should be interpreted with caution. The number of FCCHs was too small to assess the predictors of QRIS ratings, but descriptive tabulations suggest variation in mean enrollment, languages used in classrooms, funding streams, and Consortium by California QRIS rating level.

Internal consistency is low, but this is expected given the multidimensional nature of QRIS ratings.

In the context of a QRIS, internal consistency provides information about the extent to which a group of variables (in this case, element scores) produce consistent scores, thereby describing the extent to which they are unidimensional in nature, meaning that they measure a single underlying construct (in this case, program quality). If the multiple aspects of program quality measured in the QRIS are not strongly related to each other, the internal consistency will be low, indicating that the overall quality level measured by the QRIS rating is not unidimensional. If the QRIS ratings have low internal consistency, this means they are a relatively weak indicator of how a program scored on any specific element. If this is the case, the state may wish to provide element score information in addition to the overall rating so that parents can understand how programs perform on the aspects of quality that are most important to them.

California QRIS ratings are designed to measure multiple aspects of program quality and to assign a single rating reflecting the program's overall level of quality based on these multiple aspects of quality. California QRIS ratings were designed to allow a broad range of definitions of program quality, so a low to moderate level of internal consistency is to be expected. Low to moderate internal consistency does not suggest that the QRIS rating is flawed, but rather that the rating elements measure different dimensions of quality and that users of the ratings would get more reliable information about quality if they also were informed about how programs were rated on the elements.

Among the seven domains collected on centers, internal consistency is low ($\alpha = 0.54$) according to criteria used for unidimensional scales, as expected. Generally, a Cronbach's α of 0.8 or higher is considered to indicate a high level of reliability of a unidimensional scale, and between 0.7 and 0.8 is considered to be acceptable for such a scale. Internal consistency of the California QRIS ratings would increase somewhat ($\alpha = 0.67$) if the Developmental and Health Screenings Element and the Ratios and Group Sizes element were both removed from the overall rating and reported separately. In other words, an overall rating for centers based on just five of the element scores (Child Observation, Minimum Qualifications for Lead Teacher/FCCH, Effective Teacher-Child Interactions, Program Environment Rating Scale(s), and Director Qualifications) has somewhat higher internal consistency, although still not high enough to provide evidence of a single construct. Internal consistency also is relatively low among the five domains collected on FCCHs ($\alpha = 0.63$), but would not be increased by removing any element scores. The low levels of internal consistency indicate that the QRIS ratings do not measure a unidimensional program quality construct, especially among centers. In other words, the overall QRIS ratings do not represent a single dimension of quality, but rather represent diverse types of program quality.

Thus, programs with the same rating could have very different strengths and limitations. This enables a diversity of programs to achieve high ratings, providing more options for high quality, but parents choosing care will not be able to identify which programs at a specific rating level match their priorities for high-quality care (see exhibit 3.5 for an example).

These predictably low levels of internal consistency also serve as an important reminder about the likely relationships between QRIS ratings and the observed measures of program quality collected for the study analyses. As noted previously, low internal consistency across the multiple rating domains in the QRIS rubric underscores the point that the overall QRIS rating includes different elements of quality that may not be closely related to each other and are not necessarily expected to be. But because of this multidimensionality, we would not expect strong relationships between QRIS ratings and the observed quality measures or the child outcome measures collected for the study.

Exhibit 3.5. Example Element Score Variation for Two Tier 4 Centers

	Center A	Center B
Overall Rating	****	****
Elements		
Child Observation	***	*****
Developmental and Health Screenings	*****	**
Minimum Qualifications for Lead Teacher/FCCH	***	*****
Effective Teacher–Child Interactions	**	*****
Ratios and Group Size	*****	**
Program Environment Rating Scale(s)	***	*****
Director Qualifications	*****	***

The elements included in California QRIS ratings are not redundant; indeed, some pairs of elements have very low correlations.

Among centers, none of the element scores were redundant, indicating that the element scores measure different aspects of program quality. The Ratios and Group Size and Developmental and Health Screening elements had particularly low correlations with other elements (Spearman's ρ below .10 for most element pairs), as shown in exhibit 3.6, while other pairs of elements had Spearman's ρ correlations ranging from .11 to .46. Similar correlations were observed among element scores collected in FCCHs. These correlations are low, but the QRIS is designed to measure diverse aspects of quality, and programs are expected to earn different scores on rating elements. The low correlations are reflected in the relatively low internal consistency of the QRIS ratings.

Exhibit 3.6. Correlations (Spearman's ρ) Among Element Scores: Centers

	CO	DHS	MQ	CLASS	RGS	ERS	DQ
Child Observation (CO)	1.000						
Developmental and Health Screenings (DHS)	0.348*	1.000					
Minimum Qualifications for Lead Teacher/ FCCH (MQ)	0.233*	0.077	1.000				
Effective Teacher-Child Interactions: CLASS (CLASS)	0.106*	0.077	0.303*	1.000			
Ratios and Group Sizes (RGS)	–0.030	0.061	–0.128*	–0.081	1.000		
Program Environment Rating Scales (ERS)	0.195*	–0.012	0.305*	0.324*	–0.058	1.000	
Director Qualifications (DQ)	0.351*	0.051	0.464*	0.135*	–0.078	0.149*	1.000

n = 363 centers (excludes two centers serving only infants because the element score requirements are different). Correlations are calculated using Spearman's ρ, a nonparametric correlation coefficient that can be interpreted in a similar way to Pearson's r.

* $p < .05$

Elements with limited variability tend to be weakly related to the overall QRIS rating and other element scores.

Among centers, the Ratios and Group Sizes element is weakly correlated with the overall QRIS rating (Spearman's $\rho = 0.16$), while other element scores are moderately correlated with the overall QRIS rating (Spearman's ρ ranging from .45 to .57, see exhibit 3A.3 in appendix 3A). The Ratios and Group Sizes element has the least amount of variability, which may explain why it detracts from the internal consistency of the QRIS rating and has weak relationships with other elements. Among FCCHs, the Effective Teacher-Child Interactions: CLASS (CLASS) is the element most weakly correlated with the overall QRIS rating (Spearman's $\rho = 0.35$), a low to moderate correlation. This element also has limited variability among FCCHs, which contributes to its relatively weak correlation with QRIS ratings. Other element scores are more strongly correlated with the overall QRIS rating (Spearman's ρ ranging from .53 to .62).

These results indicate that some elements are not contributing as much to the overall rating. With a more diverse group of programs in the system, we might see more variability among the element scores, which might, in turn, improve the degree to which these elements are able to differentiate programs.

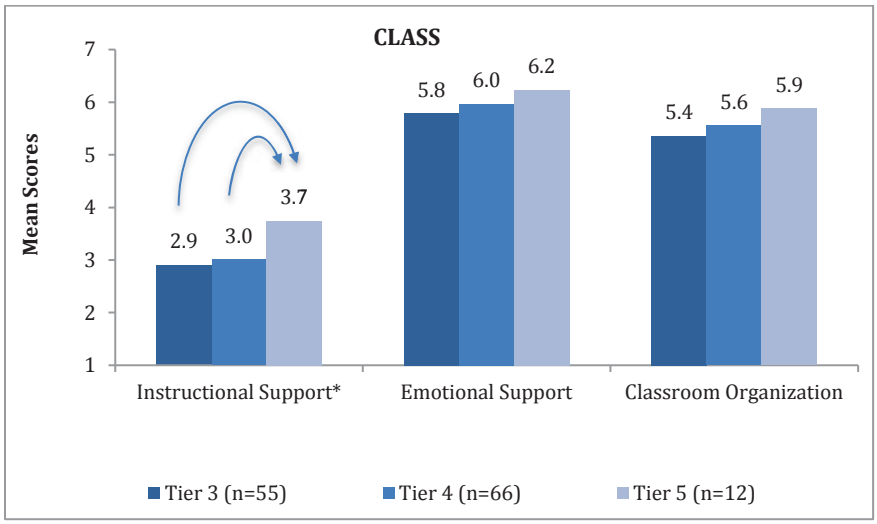
How Well Does the QRIS Differentiate Programs Based on Observed Quality?

To determine how effective the RTT-ELC rating structure is at defining and measuring quality in early learning settings, we compared the ratings assigned by the Consortia with independent measures of quality to see how closely they align. Our analyses compare QRIS rating levels and element scores from 175 fully rated centers with their scores on the independently observed measures of quality, the CLASS and the PQA. Results from these analyses find some evidence that the California QRIS ratings differentiate programs based on observed quality.

Average scores in all three Pre-K CLASS domains increase steadily as California QRIS ratings increase, but the differences are mostly small in magnitude and only the relationship with Instructional Support scores is statistically significant.

Among preschool classrooms in centers, Instructional Support scores increased steadily with each California QRIS rating level, shown in exhibit 3.7. The differences were large and statistically significant between centers rated at Tier 5 and those rated at Tier 3 (d = 0.97, close to one standard deviation in magnitude), and also between centers rated at Tier 5 and those rated at Tier 4 (d = 0.82). This magnitude of difference is meaningful, particularly because the Instructional Support domain is the most difficult one to score well on and is most strongly predictive of child cognitive skills among the CLASS domains (Howes and others 2008; Mashburn and others 2008). However, mean differences between Tiers 3 and 4 are smaller and not significant, and the majority of programs in the sample are rated at these tiers. There were no significant differences between rating levels on the Emotional Support or Classroom Organization domains, but the means on both of these domains did increase a small amount as the California QRIS rating levels increased. Exhibit 3.7 illustrates the average Pre-K CLASS domain scores by California QRIS rating level in centers and the magnitude of differences between them.

Exhibit 3.7. Average Pre-K CLASS Domain Scores by California QRIS Rating Level: Centers



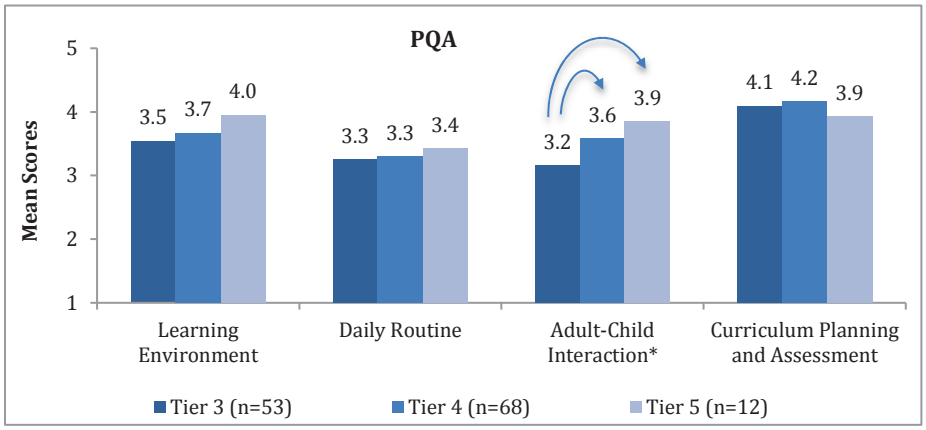
NOTE: Excludes the two centers in the Pre-K CLASS sample that were rated at Tier 2 because average CLASS score data are not reliable for rating levels with fewer than five observations.

There is a positive relationship between California QRIS ratings and three of the four PQA preschool domain scores, but only the positive relationship with Adult-Child Interaction domain scores is statistically significant.

Among centers, scores on the Adult-Child Interactions domain of the PQA increased steadily from Tier 3 through Tier 5. The differences were large and statistically significant between Tier 3 and Tier 4 (d = 0.58) and between Tier 3 and Tier 5 (d = 0.96). This domain measures many aspects of the quality of interactions between children and adults, including warmth and sensitivity, communication, child directedness, encouragement, and problem solving. The difference between Tiers 4 and 5 is smaller and not significant.

There were no significant differences in other preschool Form A subscale scores by rating level, but scores on the Learning Environment and Daily Routine domains increased consistently with tier rating. The Learning Environment domain assesses environment safety, the quality of equipment and materials, and organization of the classroom space. The Daily Routines domain assesses routines and scheduling of the day, grouping, child-directed activities, and transitions. Scores on the Curriculum Planning and Assessment domain did not increase consistent with tier rating. Exhibit 3.8 illustrates the average preschool PQA domain scores by California QRIS rating level in centers and the magnitude of differences between them.

Exhibit 3.8. Average Preschool PQA Form A Domain Scores by California QRIS Rating Level: Centers



NOTE: Excludes the one center in the preschool PQA sample that was rated at Tier 2 because average PQA score data are not reliable for rating levels with fewer than five observations.

Among the analyses examining the relationship between element scores and observed quality, only the element scores based on the CLASS and ERS consistently and significantly predict independent observation scores on the CLASS and PQA instruments (exhibit 3.26). The effective Teacher-Child Interactions element score is based on the CLASS instrument. Element scores of 1 or 2 are determined by level of familiarity with the CLASS instrument, while higher scores of 3, 4, or 5 require an independent CLASS observation (see exhibit 3.2 in chapter 3 for the Hybrid Rating Matrix). All but two centers in the analysis sample received scores of 3 or higher on this element. The effective teacher-child interaction element score positively predicts all classroom-level observation scores used in the study, including CLASS and PQA Form A total scores and subscale scores, except that the average scores on the PQA curriculum planning and assessment are slightly lower at Tier 4 than Tier 3 without statistical significance.

The Program Environment Rating Scale element is based on the ERS instrument. Element scores of 3, 4, or 5 are based on scores on an independent ERS observation, and element scores of 1 or 2 are determined by level of familiarity with the ERS instrument, although programs also may receive a score of 2 if they received an independent ERS observation but did not meet the requirement for an element score of 3. The Program Environment Rating Scale is significantly related to the total CLASS and PQA Form A scores and most of the subscale scores, but the relationships are not consistently positive. Specifically, centers with a score of 2 on the Program Environment Rating Scale element had higher average scores on some CLASS and PQA

domains than centers with a score of 3 or higher on the Program Environment Rating Scale element.

The other element scores may be thought of as indicators of structural quality, and some previous studies have found that structural quality measures predict classroom observation scores (Burchinal and others 2002; Goelman and others 2006; NICHD Early Child Care Research Network 2002; Phillips and others 2000). However, in this study of the California QRIS, none of the structural element scores positively predict average scores on the CLASS or the PQA, with the exception of a significant positive relationship between the Developmental and Health Screening element score and the PQA Parent Involvement and Family Services element score. In other cases, the structural elements were not significantly related to scores on independent measures of structural quality in the PQA instrument. This lack of a positive relationship between structural element scores and the independent structural quality measure suggests that the element scores could potentially be improved to ensure more variability. Furthermore, there were some statistically significant negative relationships—such as between the Director Qualifications element and the CLASS Classroom Organization independent observation score, and between the Child Observation element and the PQA Curriculum Planning and Assessment and also the PQA Parent Involvement and Family Services independent observation scores—that suggest that some structural elements are inconsistent with other measures of quality.

Expanding QRIS participation to a more diverse group of programs might have the effect of increasing variability in the structural element scores (which show little variability in the study sample) and improving the relationship with independent measures of structure quality. As with our findings on the distribution of California's QRIS ratings, it is important to stress that our sample of fully rated programs was small and of higher quality than programs participating in the QRIS in 2013 that had provisional ratings. Thus, results of the analyses examining the relationship between QRIS ratings or element scores and observed quality may not apply to the broader range of programs participating in the QRIS.

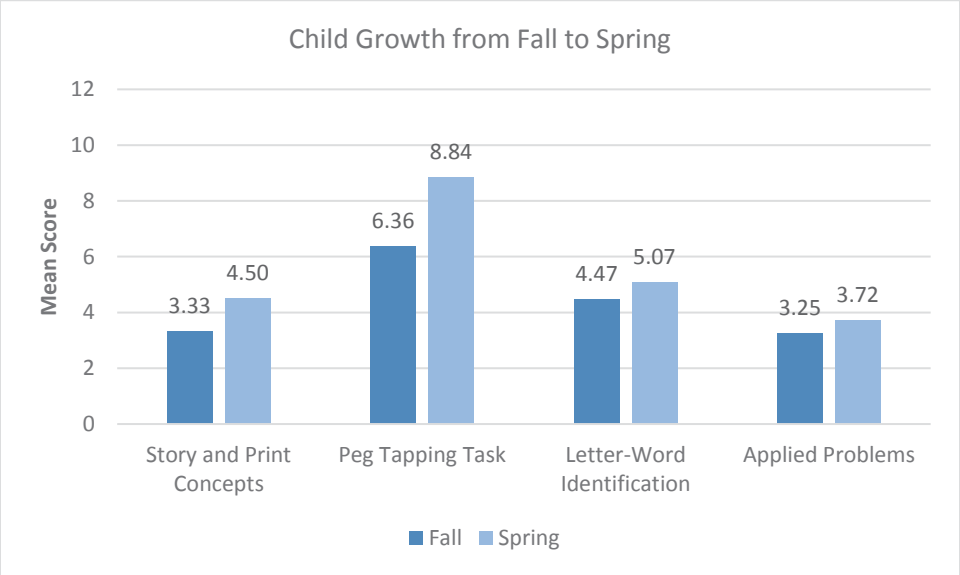
How Well Does the QRIS Predict Children's Outcomes?

To assess the effectiveness of the RTT-ELC rating structure for measuring the aspects of quality that best support children's learning and development, we compared developmental outcomes in spring 2015 for children enrolled in sites at different rating levels, while controlling for children's skills in the fall as well as child and program characteristics. These analyses examine validity of the QRIS ratings for one of many potential purposes: to identify programs in which children have the highest learning gains over the course of the year. If the ratings have strong validity for this purpose, we would expect to find higher child outcome scores, on average, for children attending higher rated programs in comparison with children attending lower rated programs. However, given the early stage of implementation and the limited variation in ratings observed, we would not expect many significant relationships, For these analyses, we draw on direct assessments of 1,611 three- and four-year-old children in 132 fully rated sites who were assessed in the fall and spring of the 2014–15 program year.

Children in the study sample demonstrate growth on developmental outcomes in literacy, mathematics, and executive function from fall to spring, though these outcomes do not increase steadily as California QRIS rating levels increase.

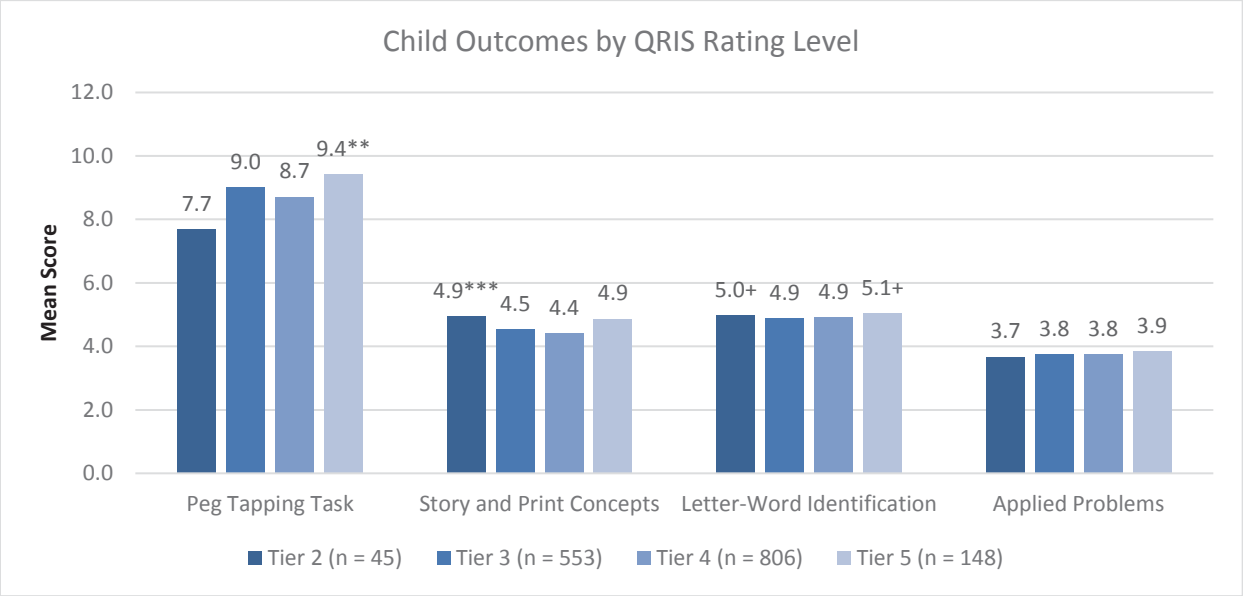
As expected, children show growth on each of the developmental outcomes measured from the beginning of the program year to the end of the programs year (exhibit 3.9). However, as shown in exhibit 3.10, we find few statistically significant differences in child outcomes for sites with different overall QRIS ratings. Compared with children in Tier 3 sites, children in Tier 5 sites had slightly higher executive function, on average, as measured by the Peg Tapping task. Mathematics skills are comparable for children across all tier levels, as are early literacy skills measured on the letter-word identification measure. In contrast, children in the five Tier 2 sites had significantly higher scores, on average, than children in the Tier 3 sites on another early literacy measure, Story and Print Concepts. In these analyses, differences between children in Tier 2 sites may not be estimated reliably because the study sample includes only five sites rated at Tier 2. The five sites rated at Tier 2 may not be representative of Tier 2 sites in general. In fact, they appear to be quite different, as they are less likely to receive standards-based funding and more likely to accept private pay and child care subsidy vouchers. They also appear to be smaller and less likely to serve children who speak Spanish at home compared with sites at other rating levels. An additional limitation of this analysis is that the study design does not allow us to account for all differences between children enrolled in programs at each rating tier. Although we controlled for program characteristics and children's home language, gender, and special-needs status, other demographics, such as family income and parent education, are not controlled for and may be influencing these relationships.

Exhibit 3.9. Raw Scores on Child Assessments for all Children, Fall and Spring



NOTE: Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

Exhibit 3.10. Adjusted Mean Scores on Child Assessments by California QRIS Rating Level: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

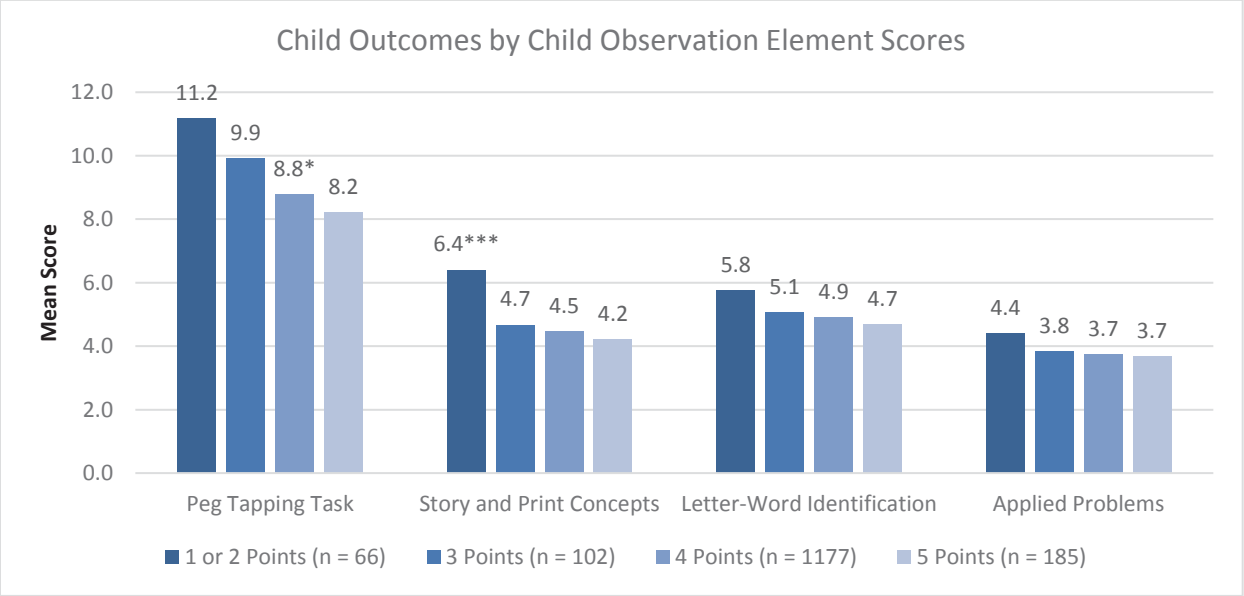
The lack of significant improvements in child outcomes by QRIS rating level contrasts with the findings of the analyses examining the relationship between ratings and observed quality, which found higher levels of classroom quality (as measured by the Pre-K CLASS Instructional Support and PQA Adult-Child Interactions domains) in higher rated programs. Although classroom interactions were significantly different among some rating levels, the differences were small in magnitude and perhaps not substantial enough to demonstrate better child outcomes.

The Child Observation element scores appear to be negatively related to child developmental outcomes.

The two elements that focus on teachers' use of screening and assessment tools—the Child Observation element and the Developmental and Health Screenings element—do not show the expected relationship to children's outcomes. In fact, the pattern of child assessment scores across Child Observation element scores suggests that children in lower rated sites are performing better (exhibit 3.11). Although the pattern is consistent across element scores and outcomes, there are the only two differences that are statistically significant. First, children in sites with four points on this element have lower executive function outcomes than children in sites with three points on this element. Second, children in sites with one or two points on this element have higher scores on the preliteracy Story and Print Concepts measure compared with those in sites with three points on this element. In the analyses examining the relationship between element scores and observed quality, the Child Observation domain was inconsistent in its relationship with other independent quality measures, including a significant negative

relationship with two domains of the PQA: Parent Involvement and Family Services, and Curriculum Planning and Assessment. However, this element score also has particularly limited variability for centers in the study sample, representing approximately three out of four children in the study sample and also about three out of four centers with fully rated sites. The analyses of this element score may not be reliable in our study sample, and results could be different in a more diverse sample of programs with a broader distribution of scores on this element.

Exhibit 3.11. Adjusted Mean Scores on Child Assessments by Child Observation Element Scores: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

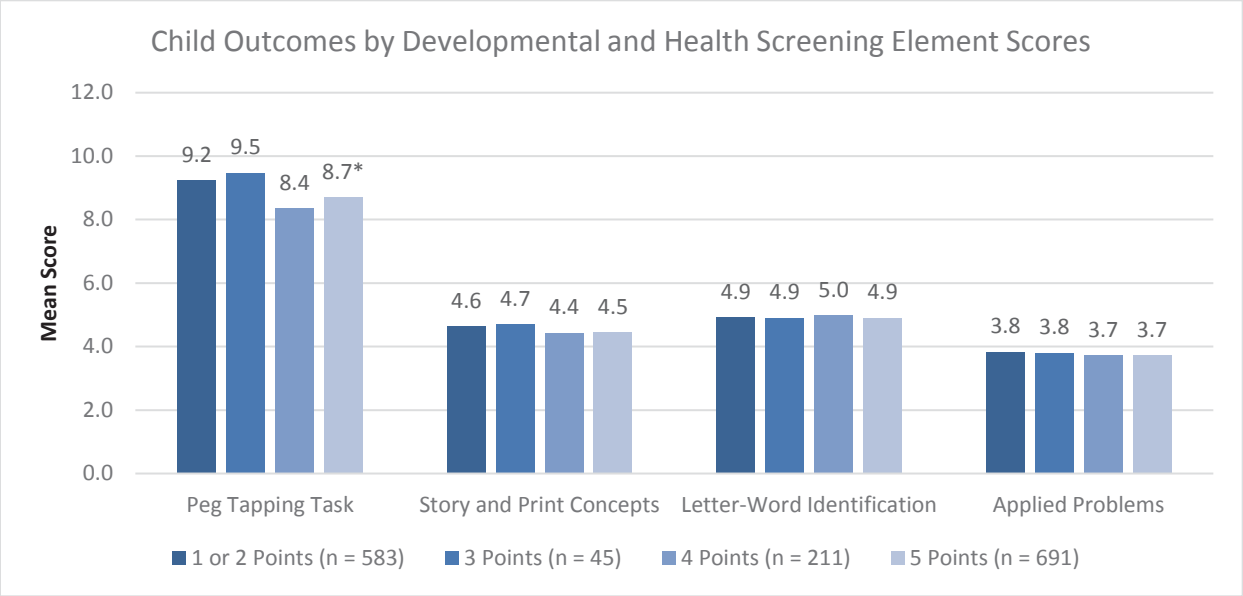
NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

The Developmental and Health Screening element is not related to children's mathematics and literacy outcomes, and appears to have a negative relationship with executive function.

Children's developmental outcomes did not increase consistently with QRIS rating level, and there were no significant differences in mathematics and literacy outcomes by rating tier. However, children in sites that received five points on the Developmental and Health Screenings element had lower scores on the executive function measure than children in sites that received three points on this element (exhibit 3.12). In addition to basic health screening, a score of 5 on this element requires the use of the Ages and Stages Questionnaire (ASQ) and ASQ-SE [Social Emotional] for child developmental screening and the use of the results for referrals and interventions in the program, while a score of 3 requires the use of any valid and reliable developmental screening tool. This element score was positively associated with the PQA Parent Involvement and Family Services score, which is an independent measure of screening and referral practices as well as parent involvement in the program. However, this element score was not associated with any other classroom observation measures in the analyses examining the relationship between element scores and observed quality, and the internal consistency analyses

found that the QRIS internal consistency would be higher without this element score. In summary, the Developmental and Health Screening element appears to be consistent with another measure of screening and referral practices, but is not consistent with other aspects of quality measured in the QRIS.

Exhibit 3.12. Adjusted Mean Scores on Child Assessments by Developmental and Health Screening Element Scores: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

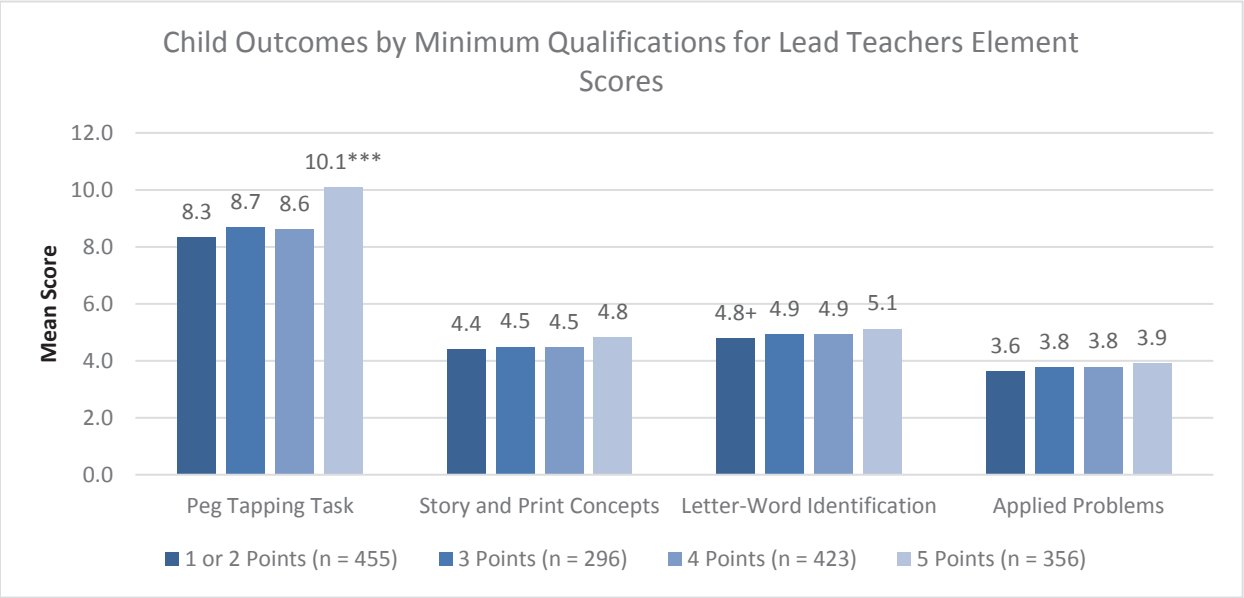
NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

The other structural quality indicators—Minimum Qualifications for Lead Teacher/ FCCH, Ratios and Group Sizes, and Director Qualifications—show some limited predictive relationship with children's outcomes.

Looking at the patterns of relationships between child outcomes and the other structural quality elements—Minimum Qualifications for Lead Teacher/FCCH (exhibit 3.13), Ratios and Group Size (exhibit 3.14), and Director Qualifications (exhibit 3.15)—we find some positive associations, but few are statistically significant. For example, children in sites receiving five points in the Minimum Qualifications for Lead Teacher/FCCH element (requiring at least a Bachelor's degree with early childhood education (ECE) units or a program director permit, plus at least 21 hours of professional development annually) have significantly higher executive function scores in comparison with those in sites receiving three points (which does not require a specific degree but does require 24 ECE units and 16 general education units or a teacher permit, plus at least 21 hours of professional development annually). The differences in child mathematics and literacy outcomes are small and nonsignificant but appear to be positive in direction for the Minimum Qualifications for Lead Teacher/FCCH element. Although we did not find statistically significant associations between this element and the quality measures used in

the analyses examining the relationship between element scores and observed quality, we do see some limited pattern of positive relationships for sites at the three-, four-, and five-point levels on the CLASS domains, and at the two-, four-, and five-point levels on the PQA domains.

Exhibit 3.13. Adjusted Mean Scores on Child Assessments by Minimum Qualifications for Lead Teachers Element Scores: Centers

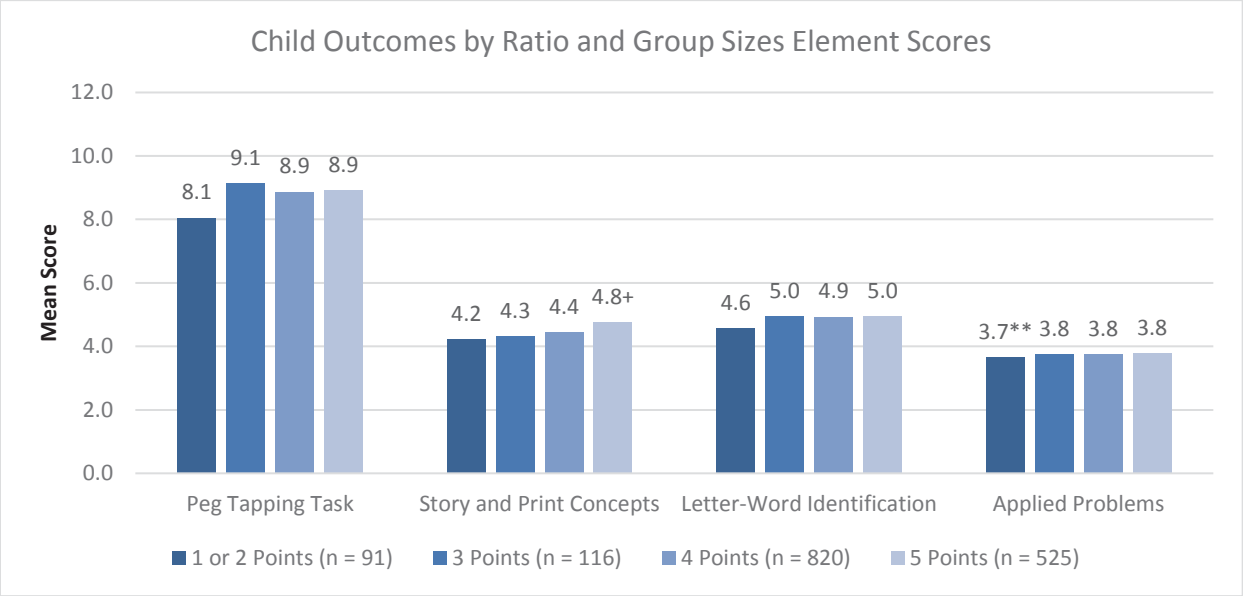


+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

There also is a positive but weak pattern for Ratios and Group Sizes element scores. Children in sites with a score of 3 on the Ratios and Group Sizes element have significantly higher scores on the Applied Problems mathematics assessment than children in sites with a score of 2, although the difference is small in magnitude (the element score requires no difference in the ratios but does have a smaller group size at Tier 3). Also, there are no differences in mathematics outcomes between children in sites with scores of at 3, 4, and 5 on the Ratios and Group Sizes element. The relationship between this element score and children's literacy and executive function outcomes is not consistent in direction and not statistically significant. This element also had inconsistent results in the concurrent validity analyses and a statistically significant negative relationship with the PQA Staff Qualifications and Staff Development observation score. Furthermore, this element had extremely limited variability, with most programs receiving a rating of 4 or 5, and internal consistency analyses suggested that the QRIS rating would have higher internal consistency without the inclusion of this element score.

Exhibit 3.14. Adjusted Mean Scores on Child Assessments by Ratios and Group Sizes Element Scores: Centers

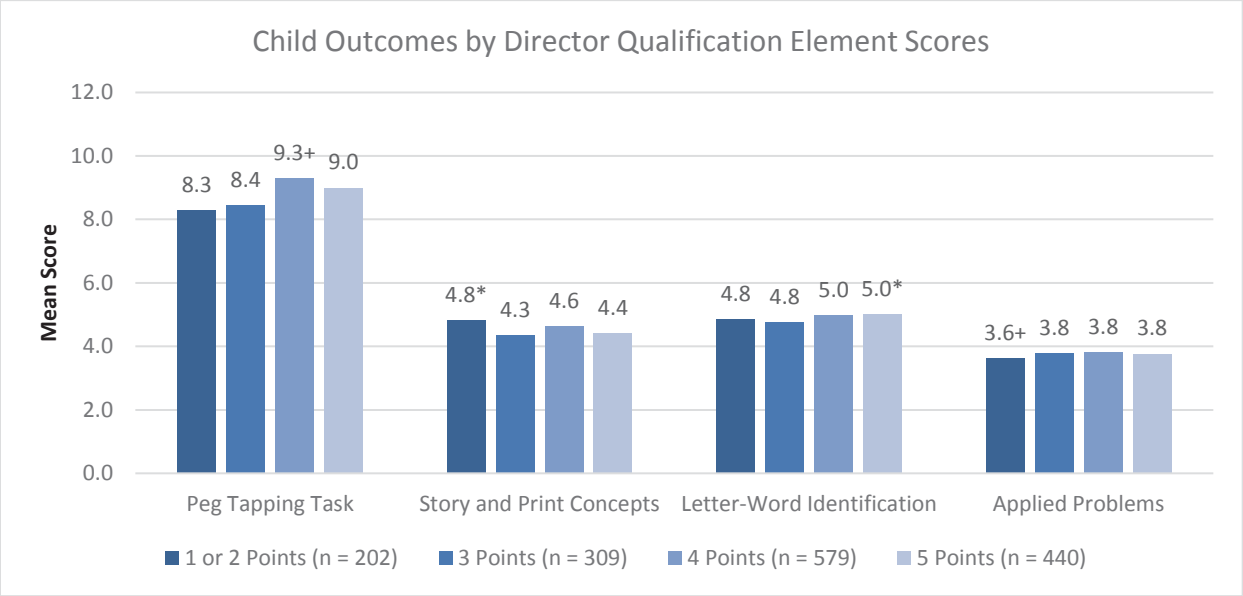


+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

For the Director Qualifications element, there is an inconsistent pattern of relationships with child literacy outcomes. Children in sites with a score of 3 on this element have significantly lower letter-word scores than children in sites with a score of 5, but significantly higher Story and Print Concepts scores than children in sites with score of 1 or 2. The relationship between Director Qualifications element scores and children's executive function outcomes is inconsistent in direction but not statistically significant, and there is no relationship between this element and children's mathematics outcomes. Relationships between this element and other observed measures of quality also were inconsistent in direction, and there was a negative relationship between the Director Qualifications element scores and the Classroom Organization observation score.

Exhibit 3.15. Adjusted Mean Scores on Child Assessments by Director Qualifications Element Scores: Centers



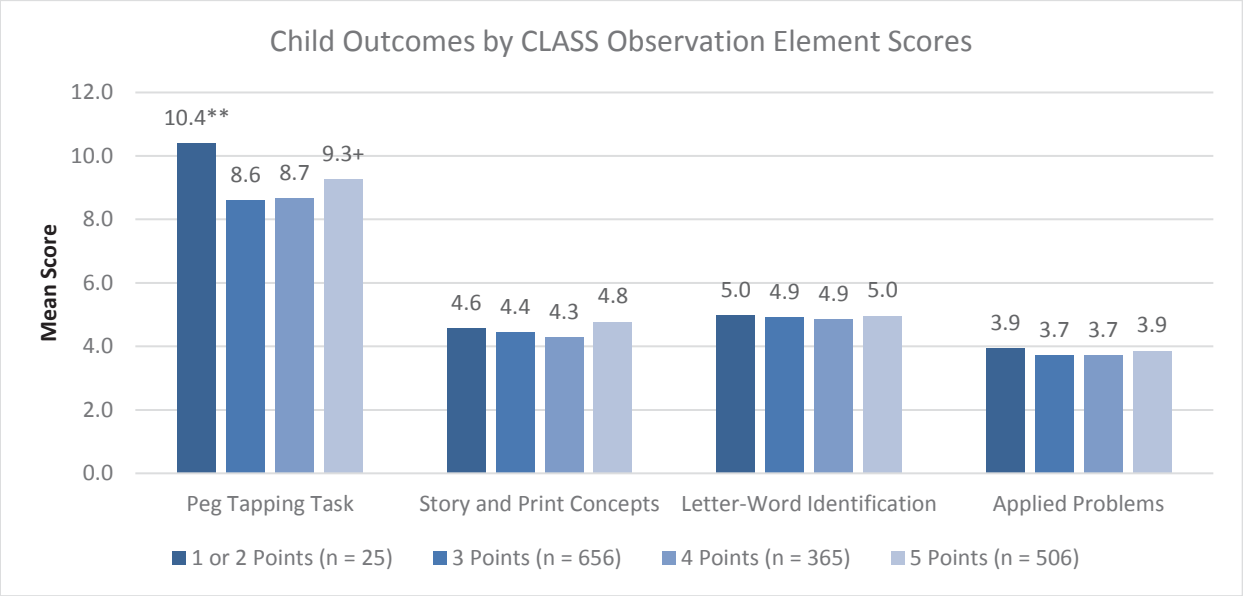
+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

The process quality elements—the CLASS Observation element and the Program Environment Rating Scale (ERS) element—do not appear to be consistently related to children's outcomes.

We might expect the elements measuring process quality—the CLASS Observation element and the Program Environment Rating Scale element—to be most predictive of children's outcomes because they are considered valid measures of program quality that are closest to the child's experience. Furthermore, both of these elements were most strongly related to independent observations of quality in analyses examining those relationships and exhibited more variability in the distribution of scores than the structural quality elements. However, there is little variation in children's outcomes across the CLASS element scores (exhibit 3.16). The only significant finding is that sites receiving five points, or one or two points, on this element have children who outperform those in sites receiving three points on the Peg Tapping task that measures executive function. However, it is important to note that at one or two points on this element, sites do not receive a CLASS observation. It is possible that these sites would have received higher CLASS element scores if they had been observed. In addition, only 3 centers with a total of 26 children in the sample received a rating of 1 or 2, so estimates for this group are less reliable than for sites that received higher scores on the CLASS element.

Exhibit 3.16. Adjusted Mean Scores on Child Assessments by CLASS Observation Element Scores: Centers

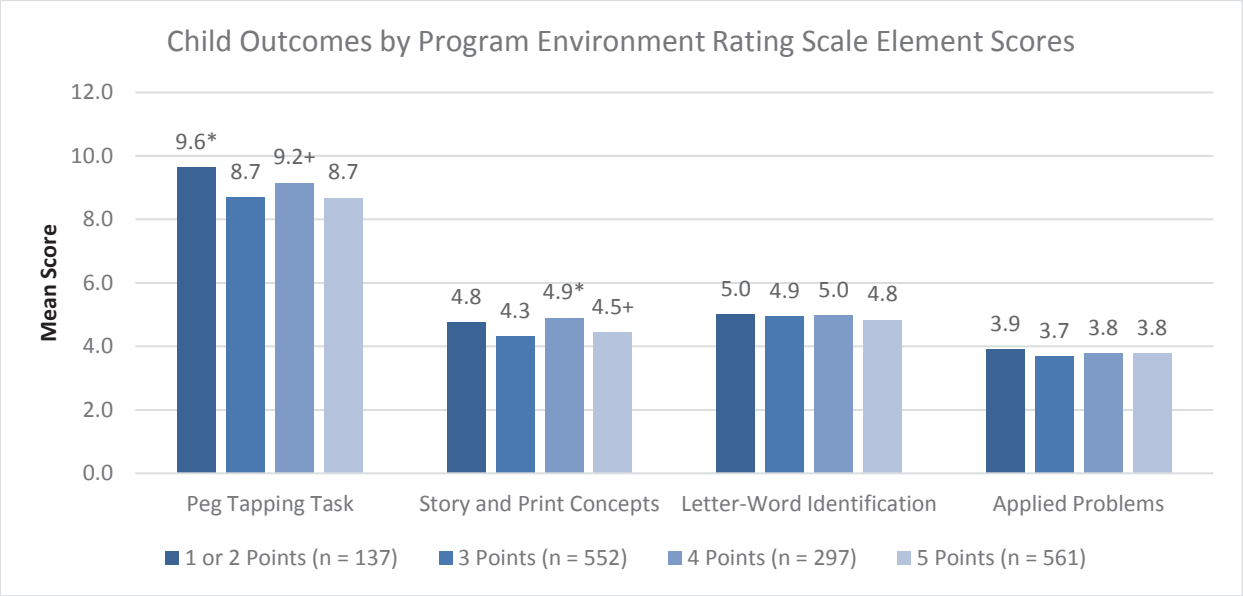


+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

Results are not consistent on the ERS element either (exhibit 3.17). Children in sites with one or two points on the ERS element perform better on the Peg Tapping task than children in sites receiving three points, although it is important to note that there are very few children in sites receiving one or two points, so this difference should be interpreted with caution. Children in sites with four points on this element also perform better on this measure, although the difference between three and four points is only marginally significant. Children in sites with four or five points on the ERS element perform better on the story and print concepts assessment than children in sites receiving three points.

Exhibit 3.17. Adjusted Mean Scores on Child Assessments by Program Environment Rating Scale Element Scores: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

In summary, the research team found limited associations between California QRIS tier ratings and children's outcomes, and thus did not find clear evidence to support the use of QRIS ratings to identify programs that best support child learning and development outcomes. This finding may be explained in part by the low internal consistency of the QRIS rating. The benefit of low internal consistency is that, for the most part, each element is measuring a distinctive facet of program quality. However, low reliability (that is, internal consistency) limits the degree to which we may observe predictive relationships with child outcomes because it is difficult to find clear, linear associations with a measure that exhibits low internal consistency. Future research is needed with a broader range of program types to determine whether low QRIS ratings (such as Tiers 1 and 2) have significantly different child outcomes in comparison with higher QRIS ratings (such as Tiers 3, 4, or 5).

We also tested associations between the individual elements of the tier ratings and children's developmental outcomes and found a mixed pattern of associations (see exhibit 3.18). However, some rating elements do seem to be more predictive of children's outcomes than others.

Exhibit 3.18. Summary of Observed Quality and Child Outcomes Analysis Results for California QRIS Rating Elements: Centers

Child Outcomes	Rating Elements: Child Observation	Develop. and Health Screenings	Minimum Qual. for Lead Teacher	Ratios and Group Sizes	Teacher-Child Interactions	Program Environment Rating Scales	Director Qualifications
CLASS (Preschool)							
Emotional Support (Preschool)	↗	⟋⟍↗	↗	⟋⟍↗	↗ *	⟋⟍↗*	⟋⟍↗
Classroom Organization (Preschool)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗	↗ *	⟋⟍↗*	⟋⟍↗*
Instructional Support (Preschool)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗	↗ *	↗ *	⟋⟍↗
PQA Form A Score (All Ages)	⟋⟍↗	⟋⟍↗	⟋⟍↗	↘	↗ *	⟋⟍↗*	⟋⟍↗
Learning Environment (Preschool)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗	↗ *	⟋⟍↗*	⟋⟍↗
Daily Routine (Preschool)	⟋⟍↗	⟋⟍↗	⟋⟍↗	↘	↗	⟋⟍↗*	↘
Adult-Child Interaction (Preschool)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗	↗ *	⟋⟍↗	⟋⟍↗
Curriculum Planning and Assessment (Preschool)	↘ *	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗*	⟋⟍↗	⟋⟍↗
PQA Form B Score (All Ages)	⟋⟍↗	⟋⟍↗*	⟋⟍↗	⟋⟍↗	↗	↗	⟋⟍↗
Parent Involvement and Family Services (All Ages)	⟋⟍↗*	⟋⟍↗*	⟋⟍↗	⟋⟍↗	↗	↗	↘
Staff Qualifications and Staff Development (All Ages)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗*	↗	⟋⟍↗	⟋⟍↗
Program Management (All Ages)	⟋⟍↗	⟋⟍↗	⟋⟍↗	⟋⟍↗	↗	↗	⟋⟍↗

http://www.cde.ca.gov/sp/cd/rt/aavsummaryanalyses.asp

Child Outcomes	Rating Elements: Child Observation	Develop. and Health Screenings	Minimum Qual. for Lead Teacher	Ratios and Group Sizes	Teacher-Child Interactions	Program Environment Rating Scales	Director Qualifications
Child Assessments							
Peg Tapping Task	↘ *	∧↗ *	↗ *	↗	∧↗ *	∧↗ *	↗
Story and Print Concepts	↘ *	∧↗	↗	↗	∧↗	∧↗ *	∧↗ *
Woodcock-Johnson Letter-Word Identification	↘	∧↗	↗	↗	∧↗	∧↗	↗ *
Woodcock-Johnson Applied Problems	↘	∧↗	↗	↗ *	∧↗	∧↗	↗

NOTE: Each row references the results of a separate ANOVA model.

***** indicates a statistically significant relationship, and the arrows indicate the direction of the relationship between QRIS ratings and observed classroom quality scores, for rating levels with more than five observations:

↗ indicates a consistently positive relationship; ↘ indicates a consistently negative relationship; ∧↗ indicates relationships that are not consistent in direction.

http://www.cde.ca.gov/sp/cd/rt/aavsummaryanalyses.asp

How Do Alternative Rating Approaches Affect the Distribution and Validity of Ratings?

In addition to examining how the rating and elements of the rating are related to other measures of quality and children's developmental outcomes, we also examine the rating structure. Ratings for a QRIS can be calculated many different ways. One way is California's hybrid rating method, which combines a block approach at Tier 1 with a points-based approach at Tiers 2 through 5. Other states use points at every tier, while still others use a block approach for every tier. To explore whether a different rating approach might more effectively differentiate programs on other measures of quality or might better predict children's developmental outcomes, we tested three alternative rating approaches using the same element scores collected for the California QRIS ratings (shown in exhibit 3.19). The state currently allows one of the alternative approaches (two-level block) as a local adaptation to the statewide rating approach. The five-level block approach is used in other states, and the element average approach was selected as a simplified approach to calculating the ratings.

We recalculated ratings using each of these approaches, then examined the distribution of ratings that resulted from each approach to see how it was affected by the new method. We also examined the relationship between each alternative rating approach and observed quality and child outcomes to see how they compare with the validity results for California's hybrid method for calculating the ratings.

Exhibit 3.19. Alternative Rating Approaches Examined in This Study

Rating Type	Rating Definition
California QRIS	Tier 1 is blocked; Tiers 2–5 are point-based for programs meeting block criteria for Tier 1: Rating is determined by total points earned across elements. As noted above, local Consortia have the autonomy to make some modifications to the rating structure. This is California's rating approach without local adaptations to the way the ratings are calculated using the element scores.
Two-Level Block	Tiers 1 and 2 are blocked, and Tiers 3–5 are point-based for programs meeting block criteria for Tier 2. Some Consortia have revised California's rating approach in this way.
Five-Level Block	Tiers 1–5 are blocked.
Element Average	Scores are determined by taking the average of all applicable rating elements. Averages are rounded to whole numbers (round up for 0.5 and greater, round down for less than 0.5).

The distribution of ratings varies by rating approach, especially when blocks are used.

First, we found that the distribution of rating levels varies by rating approach (see exhibit 3.20 for centers and exhibit 3.21 for FCCHs). The largest changes occur in rating approaches using blocks; 22 percent of centers and 63 percent of FCCHs have lower ratings when only Tiers 1 and 2 are blocked, while 93 percent of centers and 94 percent of FCCHs would be assigned lower ratings when all five tiers are blocked (see exhibit 3.22).

Exhibit 3.20. Distribution of Ratings Using Alternative Rating Approaches: Centers

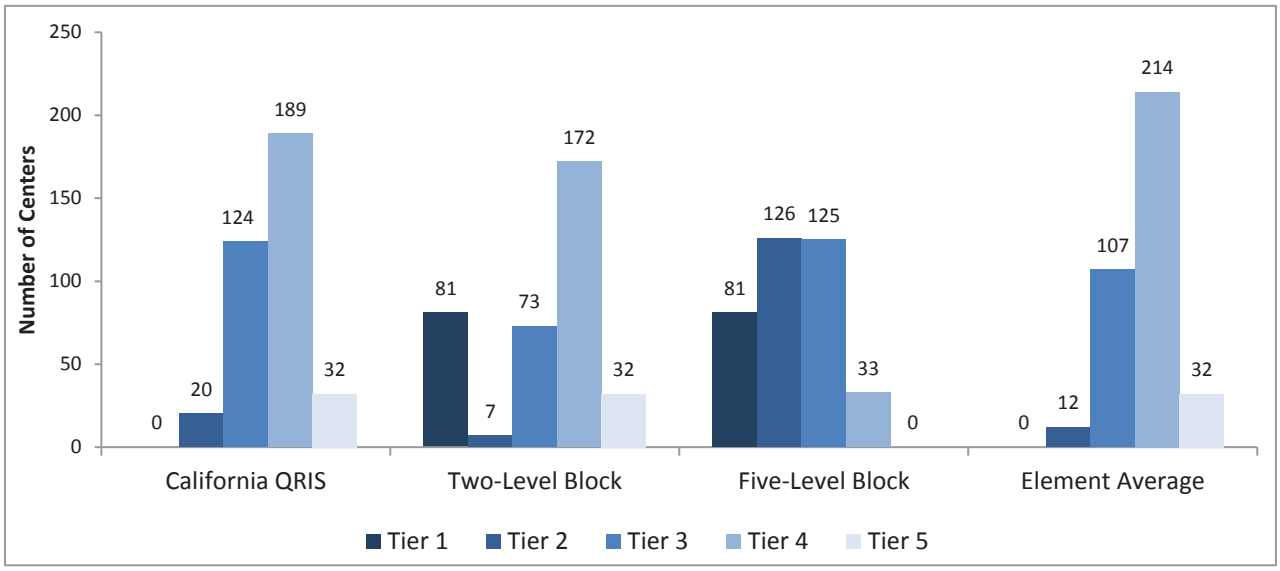


Exhibit 3.21. Distribution of Ratings Using Alternative Rating Approaches: FCCHs

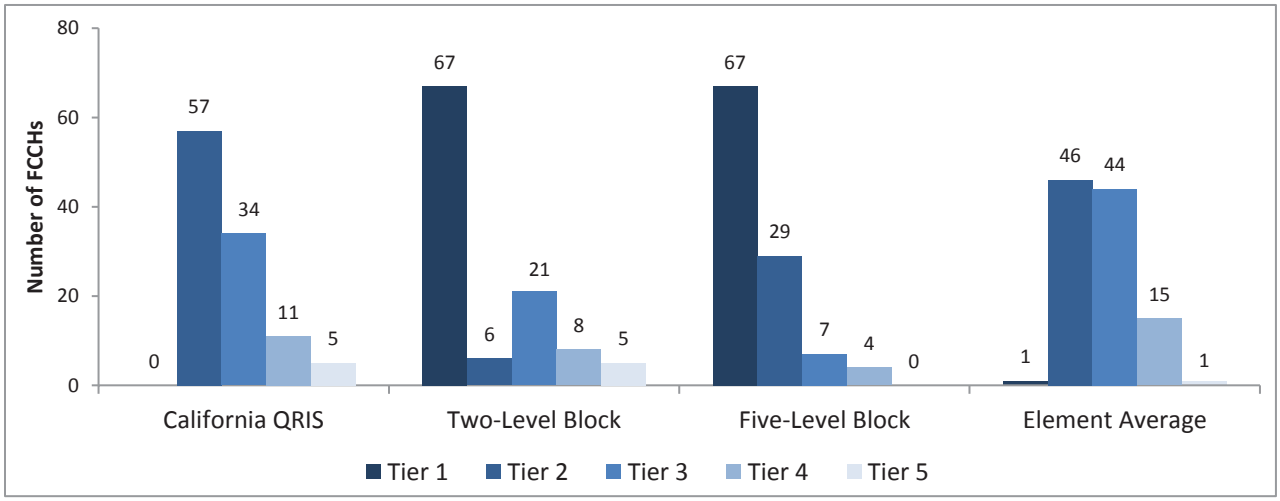


Exhibit 3.22. Reclassification Rates for Alternative Rating Approaches: Centers

Rating Type	Percentage of Programs Rated Lower Than California QRIS Rating	Percentage Same as California QRIS Rating	Percentage of Programs Rated Higher Than California QRIS Rating
Centers (*N* = 365)			
Two-Level Block	22.2	77.8	0.0
Five-Level Block	92.6	7.4	0.0
Element Average	0.0	91.0	9.0
FCCHs (*N* = 107)			
Two-Level Block	62.6	37.4	0.0
Five-Level Block	94.4	5.6	0.0
Element Average	4.7	86.0	9.4

Element average ratings are more effective than California QRIS ratings at differentiating centers by CLASS and PQA classroom observation scores.

Ratings derived by averaging element scores—or "element average ratings"—have statistically significant positive relationships with CLASS total scores and all three Pre-K CLASS domain scores, while the California QRIS ratings are only significantly related to Instructional Support scores. Element average ratings are positively associated with the Learning Environment domain of the preschool PQA, as well as the Adult-Child Interaction domain, and relationships with the other PQA observation scores are positive in direction, although not statistically significant. Unlike the California QRIS ratings, the direction of the relationship between element average ratings and PQA program-level Form B scores also is mostly positive, although not statistically significant.

Ratings using blocks are less effective than California QRIS ratings at differentiating centers by CLASS scores, but five-level blocks are more effective at differentiating centers according to the PQA observation scores.

Exhibit 3.26 shows that rating approaches using blocking are not positively related to CLASS domain scores in most cases, in contrast to California QRIS ratings, which were related to CLASS Instructional Support scores. The relationship with CLASS scores is weakest in ratings that block at all five tiers, and CLASS scores do not consistently increase as the rating level increases. However, ratings with blocking at all five rating levels are more predictive of PQA classroom observation scores than California QRIS ratings. The five-level block ratings are positively associated with PQA Form A total scores as well as the preschool adult-child interaction domain score and are significantly related to the preschool learning environment domain score, although the relationship is not consistently positive.

When interpreting the analyses examining the relationship between alternative rating approaches and observed quality, it is important to remember that the results are specific to the sample of centers with full ratings that are included in the study, and that the relationships between alternative rating approaches and observed quality scores may differ for a more diverse set of programs in California.

How Well Do Alternative Rating Approaches Predict Children's Outcomes?

To evaluate whether alternative ratings approaches better predict children's outcomes, the study team repeated analyses testing associations between tier rating levels and children's outcomes for four approaches to calculating the ratings: California's QRIS rating approach, the element average approach, the two-level block approach, and the five-level block approach.[8]

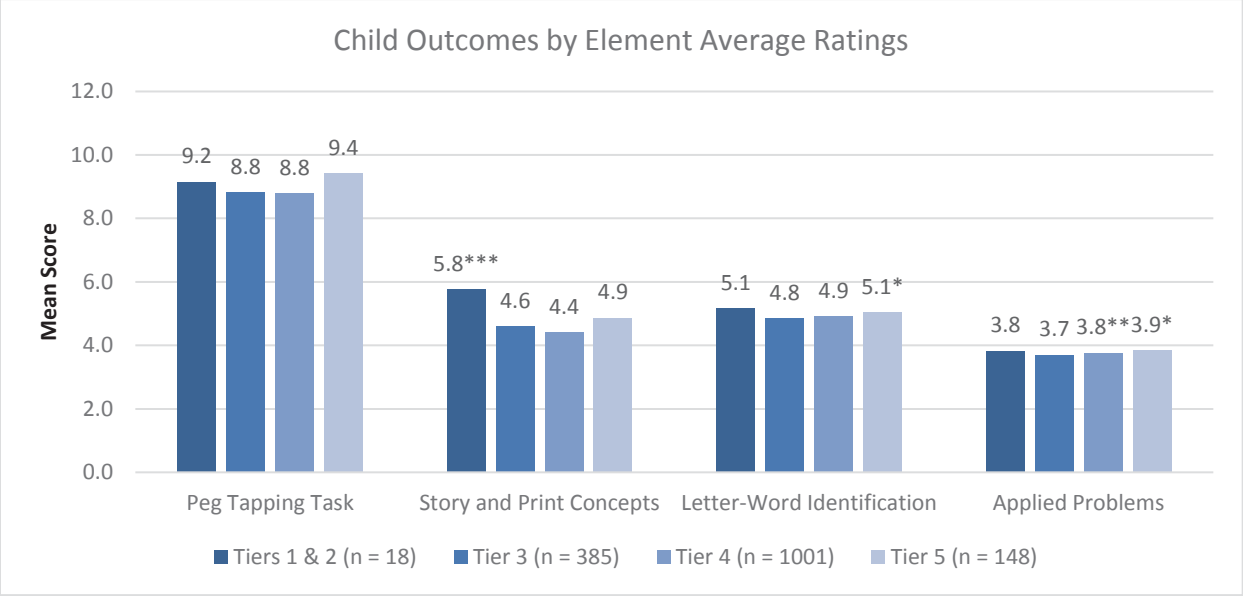
Using an element average rating exhibits somewhat better predictive relationships with child outcomes, as compared with the California QRIS rating, while two-level and five-level blocks do not.

Using the element average score appears to be somewhat more effective at predicting children's developmental outcomes compared with the California QRIS ratings (exhibit 3.23). Although the significant difference between rating levels on the Peg Tapping task measures disappears when we move from the California QRIS rating approach to the element average approach, differences on the Letter-Word Identification and Applied Problems subtests, although still small, become statistically significant, indicating a more consistent relationship using the element average approach. The patterns of relationships look quite similar for the ratings using two- and five-level blocks, though we find fewer statistically significant relationships.

Blocking scores at Tiers 1 and 2 also resulted in a similar pattern of association between rating levels and children's outcomes compared with the California QRIS ratings (exhibit 3.24). Children in centers with a tier rating of 1 or 2 combined or Tier 5 performed better on the Peg Tapping task than children in centers with a rating of 3. Results for Letter-Word Identification indicate that children in centers with ratings of 4 or 5 scored better than children in sites with a rating of 3. No significant association emerged between rating levels with a two-level block and Book and Print Familiarity or the Applied Problems assessment of early mathematics skills.

[8] We also tested the association between the total number of points earned in the rating process prior to being converted to a tier rating based on cut scores and children's outcomes. The total number of points on the rating components is a continuous measure of quality. No statistically significant association emerged between total points and children's executive function, literacy, or mathematics skills.

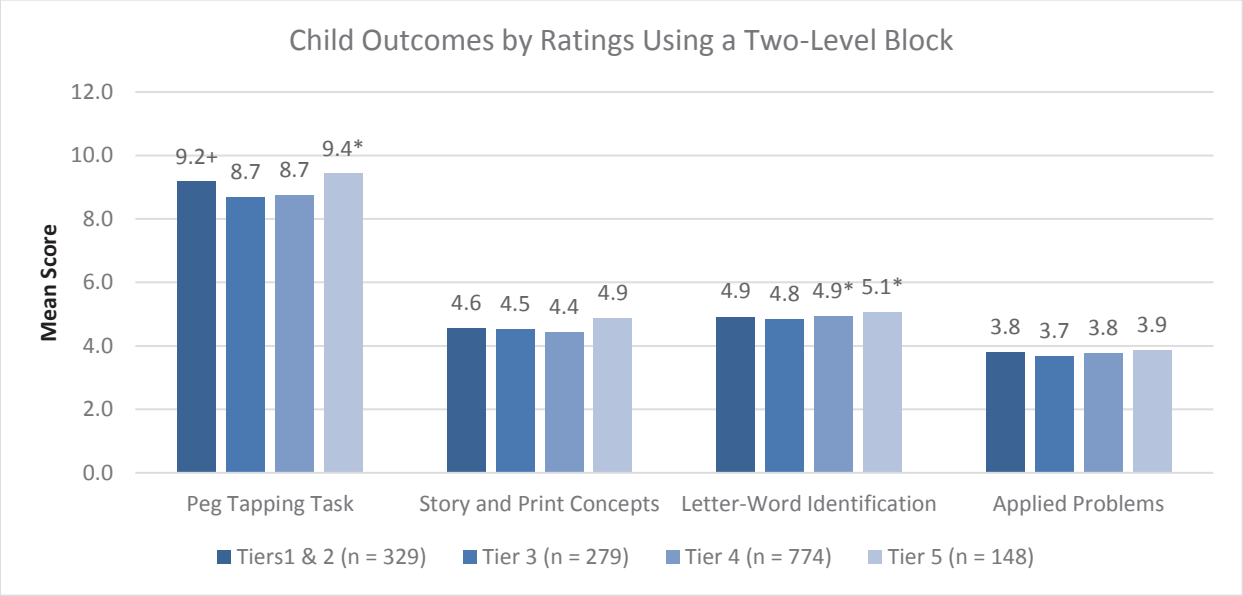
Exhibit 3.23. Adjusted Mean Scores on Child Assessments by QRIS Ratings Using the Element Average Rating Approach: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

Exhibit 3.24. Adjusted Mean Scores on Child Assessments by QRIS Rating Using the Two-Level Block Approach: Centers

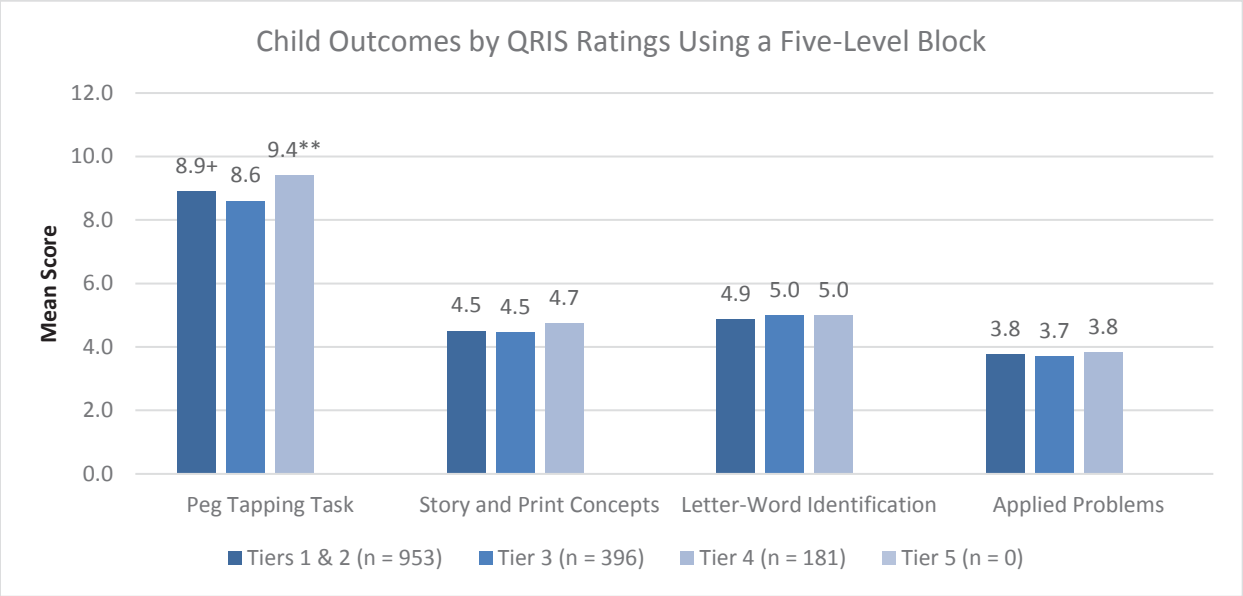


+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

Using a five-level block approach truncates the range of tier ratings in the analytic sample for the child outcomes study, such that no centers held a ratings of 5 (exhibit 3.25). Across these rating levels, statistically significant differences exist for only one of the outcomes. On the Peg Tapping task, the measure of executive function, children exhibited higher regression-adjusted scores at Tiers 1 and 2 combined and 4 than they did at Tier 3. For the remaining three measures of early literacy and mathematics skills, we found no relationship between rating levels and children's learning and skills.

Exhibit 3.25. Adjusted Mean Scores on Child Assessments by QRIS Ratings Using the Five-Level Block Approach: Centers



+ $p < .10$, * $p < .05$; ** $p < .01$.

NOTE: Sites with a rating of 3 were the reference category. Scores should not be compared across assessments; age-equivalent scores are presented for Letter-Word Identification and Applied Problems, and raw scores are presented for Peg Tapping and Story and Print Concepts.

Exhibit 3.26 provides an overview of the observed quality and child outcomes analyses using the alternative rating approaches.

Exhibit 3.26. Summary of Observed Quality and Child Outcomes Analysis Results for California QRIS Ratings and Alternative QRIS Rating Approaches: Centers

	California QRIS Rating	Two-Level Block	Five-Level Block	Element Average
CLASS Scores				
Total Score (Preschool and Toddler)	↗ *	∧↗ *	∧↗	↗ *
Emotional Support (Preschool)	↗	∧↗	∧↗	↗ *
Classroom Organization (Preschool)	↗	↗	∧↗	↗ *
Instructional Support (Preschool)	↗ *	∧↗ *	∧↗	↗ *
PQA Scores				
Form A Score (All Ages)	↗	↗	↗ *	↗
Learning Environment (Preschool)	↗	∧↗	∧↗ *	↗ *
Daily Routine (Preschool)	↗	↗	↗	↗
Adult-Child Interaction (Preschool)	↗ *	↗ *	↗ *	↗ *
Curriculum Planning and Assessment (Preschool)	∧↗	∧↗	∧↗	↗
Form B Score (All Ages)	∧↗	∧↗	∧↗	↗
Parent Involvement and Family Services (All Ages)	∧↗	∧↗	↗	∧↗
Staff Qualifications and Staff Development (All Ages)	∧↗	∧↗	∧↗ *	↗
Program Management (All Ages)	∧↗	∧↗	∧↗	↗
Child Outcomes				
Peg Tapping Task	↗ *	∧↗	∧↗ *	∧↗
Story and Print Concepts	∧↗ *	∧↗	∧↗	∧↗ *
Woodcock-Johnson Letter-Word Identification	∧↗	↗ *	∧↗	∧↗ *
Woodcock-Johnson Applied Problems	∧↗	∧↗	∧↗	↗ *

NOTE: Each row references the results of a separate ANOVA model.

* indicates a statistically significant relationship, and the arrows indicate the direction of the relationship between QRIS ratings and observed classroom quality scores for rating levelstier with more than five observations:

↗ indicates a consistently positive relationship; ↘ indicates a consistently negative relationship; ∧↗ indicates relationships that are not consistent in direction.

http://www.cde.ca.gov/sp/cd/rt/aavsummaryanalyses.asp

Summary

In this chapter, we examined the validity of the QRIS ratings for the different purposes for which the ratings could be used. These purposes include: to serve as reliable and meaningful ratings to inform parents about program quality, to differentiate programs according to the quality of program structures and adult-child interactions, to inform program quality improvement efforts, and to identify programs that best support child learning and developmental outcomes. The validity analyses for the California QRIS ratings included examining the measurement properties of the ratings, the relationship between ratings and observed quality, and predictive relationships with child outcomes. We also examined alternative approaches to calculating the ratings and examined some aspects of the validity of these ratings.

Measurement Properties

Analyses reveal that the distribution of ratings is limited and does not span all five possible QRIS rating levels. Moreover, the distribution of ratings is very different for centers and FCCHs, and is even more limited within each program type. This truncation may be due at least in part to the population of programs participating in the system: given that participation is voluntary, programs that might score lower have little motivation to become involved. In fact, many of the programs with full QRIS ratings are California Title 5 state-funded programs (State Preschool, General Child Care, or Cal-SAFE) or CSP sites; CSP funding and Title 5 funding are statistically significant predictors of California QRIS rating level among centers. The limited distribution means that the full range of ratings cannot be fully evaluated.

Internal consistency analyses find weak associations among the different rating elements. This low internal consistency does not necessarily suggest that the rating data are flawed, but rather that the aspects of quality measured in the different elements are not always closely related to each other. Elements that produce limited score variability (such as Ratios and Group Sizes for centers and Effective Teacher-Child Interactions [CLASS] for FCCHs) are weakly related to QRIS ratings and to other element scores. The evidence suggests that a program's overall California QRIS rating (which is based on combining the element scores) does not represent a single type of quality, but rather includes diverse types of programs that reach high quality ratings in different ways. Thus, the evidence suggests that the overall rating on its own may not provide parents with sufficient information about the specific aspects of program quality in which they may be interested. Providing parents with the element scores in addition to the overall rating level will better serve this purpose of the QRIS.

Evidence of Validity for the Overall Rating

Results from the analyses examining the relationship between QRIS ratings and observed quality provide some evidence that the California QRIS ratings differentiate programs according to the quality of adult-child interactions, although the differences are small in most cases. In particular, California QRIS ratings positively and significantly predict, Preschool CLASS Instructional Support scores, and Preschool PQA Adult-Child Interaction scores (see exhibit 3.26 for a summary of results). The results provide preliminary evidence supporting use of the ratings to differentiate programs according to the quality of adult-child interactions, but differences between Tiers 3, 4, and 5 are small and further research is needed to determine whether there is

better differentiation comparing programs at lower rating tiers (Tiers 1 and 2) and higher tiers (Tiers 3, 4, and 5).

Results from analyses examining children's developmental outcomes by rating tier provide limited evidence of predictive relationships. We find few statistically significant relationships between tier rating and measures of children's developmental outcomes that were included in the study. Given that we did not find consistent positive associations between ratings and child outcomes, the study does not provide evidence supporting the validity of the ratings for the purpose of identifying programs that best support child outcomes. Like other findings in this report, however, this result is not necessarily conclusive. The lack of association between California QRIS tier ratings and children's outcomes may be explained, at least in part, by the low internal consistency of the QRIS rating because it is difficult to find any clear, linear associations with a measure that exhibits low internal consistency. Limited variability in the program sites included in the early implementation of the QRIS may play a role in both the low internal consistency and the lack of relationships with child outcomes, and results could differ with a broader range of programs. In addition, the ratings could be associated with other child outcomes that were not measured in this study.

Evidence for the Validity of the Components of the Rating

Looking more closely at the elements that comprise the overall rating, there are several interesting findings to note. First, perhaps the best validity evidence is for the Effective Teacher-Child Interactions element, where we find positive relationships that also are statistically significant for each of the CLASS domains and for three domains of the PQA—Adult-Child Interactions, Learning Environment, and Curriculum Planning and Assessment—such that sites with higher element scores also have higher scores on these measures of observed quality. In terms of children's outcomes, if we exclude the three sites that received one or two points, the trends are generally positive, with higher assessment scores for children in sites with higher scores on this element; however, none of these relationships is statistically significant. In addition, the pattern looks different when we include the sites with two points, which tend to have children who score higher than the sites with three points on this element. It is important to remember, though, that sites receive two points on this element for having one person on staff who is familiar with the CLASS tool; they do not have to have an independent CLASS observation conducted. Thus, these sites are not actually evaluated on their teacher-child interactions. Also, although simply being assessed would increase these sites' score to a 3, it is possible (and even likely, given the CLASS data that we collected on these sites) that their score on this element would be even higher if they were to receive a CLASS observation. Nevertheless, the CLASS element appears to be doing a good job of differentiating programs based on quality, at least for those that receive a CLASS observation for this element.

Second, we found some evidence for the validity of the Minimum Qualifications for Lead Teachers/FCCH element. When we look at the child outcome data, we see a positive pattern of relationships, such that children in sites receiving more points on this element have higher assessment scores. Only one of these comparisons is statistically significant; children in sites at the five-point level (where teachers have a Bachelor's degree) outperform children in sites at the three-point level (where teachers have 24 ECE units plus 16 general education units or a teacher permit). We do not find any statistically significant relationships between the Minimum

Qualifications element and our quality measures. However, we do see some patterns across our measures. For example, when we consider the CLASS domains, we see a positive (though not statistically significant) relationship between the Minimum Qualifications element and scores on each of the CLASS domains for sites receiving three, four, or five points. Sites that received two points on this element (those with 24 ECE units or an associate teacher permit) do not fit the pattern; they appear to score higher than sites receiving three points, although, again, this difference is not statistically significant. Similarly, for the PQA, if we ignore sites at the three-point level, we see a positive relationship with PQA scores for sites receiving two, four, and five points, although, again, the differences are not statistically significant. This suggests that adjusting the cut points on this element could potentially improve its validity.

Third, there is the least evidence supporting the validity of the Child Observation element. There are four statistically significant relationships across all of our measures, and they are all negative. On two quality measures—the Curriculum Planning and Assessment domain of the PQA and the Parent Involvement and Family Services domain of the PQA Form B—sites with four points on the Child Observation element score higher than sites with five points. This pattern of sites at the four-point level outperforming sites at the five- point level on quality measures is fairly consistent, although only these two PQA domains reach statistically significant levels. There are relatively few sites with one, two, or three points on this element, which limits our ability to find other statistically significant relationships. We might expect that these two domains of the PQA, which specifically address child assessment and communicating with families about child assessment, to be positively related to this element. The feature that distinguishes the five-point level from the four-point level, however, is a specific aspect of practice related to assessment: the use of the Desired Results Developmental Profile (DRDP) Tech and use of its reports to inform curriculum planning. To receive four points, staff must use the DRDP twice a year to inform curriculum planning; to receive five points, staff also must upload their data into DRDP Tech. It may be that the use of DRDP Tech is not helping teachers to better use or share the assessment data, or it might be that sites that can afford to use this tool (given the technology infrastructure needed) are different in other ways that affect their curriculum planning, assessment, and family involvement practices. The pattern of relationships with child assessment scores is fairly consistent, with lower scores at higher point levels on the Child Observation element. There are two significant relationships, though, with children in sites at the two-point level outscoring children in sites at the three-point level on Story and Print Concepts, and children in sites at the three-point level outscoring children at the four-point level on the Peg Tapping task. Given the definition of the point levels on this element and that the use of the DRDP is so tied to its funding source, this element does not appear to be successfully differentiating programs based on quality and children's outcomes.

Given the variability in results across the different elements, these findings suggest that the most meaningful information about quality may come from the element scores rather than the overall rating, and Consortia may wish to provide the element scores to parents in addition to the overall rating.

Alternative Rating Approaches

We also examined how ratings would change under different rating calculation approaches. First, we found that the distribution of rating levels varies substantially by rating approach. The largest

changes in the distribution of ratings occur when ratings rely on block designs, in comparison with the California QRIS rating approach. Second, we found that element average ratings are more effective than California QRIS ratings at differentiating centers by CLASS and PQA classroom observation scores. The element average ratings also do a somewhat better job of predicting children's developmental outcomes, in particular in terms of children's literacy and mathematics skills.

In addition, although ratings using blocks are less effective than California QRIS ratings at differentiating centers by CLASS scores, five-level blocks are more effective than California QRIS ratings at differentiating centers according to the PQA observation scores. Ratings using blocks—both two- and five-level blocks—were not dramatically different from the element average rating at predicting children's outcomes, although fewer of the relationships were statistically significant.

Although some evidence supports the validity of the Hybrid Rating Matrix in its current form, an element average rating approach appears to hold the most promise from among the alternative rating approaches tested. Ratings calculated by taking an average score across elements are more effective than the California QRIS ratings at differentiating centers by CLASS and PQA classroom observation scores. They also are somewhat more effective at predicting children's literacy and mathematics skill development.

It is important to remember when interpreting these analyses using alternative rating approaches that they are specific to the sample of centers included in the study. As noted previously, the sample of programs is small and not representative of the entire population of programs in California, and the relationships between alternative rating approaches and observed quality scores may differ for a more diverse group of programs in California.

Chapter 4. Perspectives on Quality Elements and Ratings

Key Findings

This chapter describes stakeholder views of the ratings, rating elements, and the potential for publicizing ratings to a broad audience. We draw on interviews with early learning and care providers and focus groups with parents conducted in the spring and summer of 2015.

- Providers and parents generally agreed that the QRIS rating elements included the important aspects of quality.
- Parents indicated an interest in having access to the QRIS rating information, and, for the most part, wanted detailed rating information rather than a single summary rating.
- Parents currently rely on a variety of sources to inform their choices about early learning programs, including recommendations from family and friends, and online resources.
- Parents described comfort with the site and staff, convenience of the site to their home or work, program schedule, and cost of the care as key factors in their decisions about selecting a program for their child.
- Providers and parents also discussed additional quality factors—beyond those delineated in the rating matrix—that influenced their perceptions of child care, such as family engagement, the importance of the child's positive experiences at the site as well as development of school readiness skills, and a good curriculum.

While chapter 3 provided evidence concerning the validity of the rating system, this chapter describes stakeholder views of the ratings, rating elements, and the potential for publicizing ratings to a broad audience. We draw on interviews with early learning and care providers—including center teachers and directors and family child care home (FCCH) providers—in the 11 focal Consortia, and focus groups with parents in all 17 Consortia conducted in the spring and summer of 2015.

To gather their input on the ratings and rating elements, each of the providers was e-mailed a copy of the Hybrid Rating Matrix and asked to comment on whether the Quality Rating and Improvement System (QRIS) was measuring the right aspects of quality, which areas were difficult to attain, and whether any other aspects of quality should be included. In focus groups, parents were read simplified statements reflecting each of the rating elements and asked to share their awareness of each component and their perspectives about the importance of each. Parents in focus groups also weighed in on other elements, beyond those in the rating matrix, that influence their perceptions of child care quality and their care choices.

Perspectives on Elements of the Hybrid Rating Matrix

Providers and parents generally agreed that the rating matrix included the important aspects of quality.

Providers and parents alike agreed that the rating matrix is generally measuring the right aspects of quality. Though they did not necessarily agree with how all of the elements were specified, they did not think that any element should be eliminated. As one FCCH provider explained, "The rating matrix represents the right areas.... I wouldn't change any of them or add any others." Another classroom teacher described, "Overall, it's assessing the program and the children and how effective it is for their learning. All of the components are important." Similar to providers, one parent said, "They [the elements] go hand in hand, they're hard to separate." Another parent agreed that "all of them are equally important."

Analysis Approaches

- Interview and focus group transcripts were analyzed using qualitative data analysis techniques to identify common themes and response patterns.

Data and Sample

- Interviews were conducted with 25 early learning and care providers, including 13 center directors, 7 FCCH owners, and 5 center teachers across the 11 focal Consortia.
- A total of 17 focus groups—1 in each Consortium—were conducted with parents who had their children in a care setting participating in the QRIS; in total, 146 parents participated.

Core I: Child Development and School Readiness

During interviews and focus groups, providers and parents shared feedback on *Core I: Child Development*, the first of three domains that comprise the Hybrid Rating Matrix. *Core I* includes two elements: (1) *Child Observation*, which focuses on the teachers' use of observation tools to monitor children's development and learning, such as the Desired Results Developmental Profile (DRDP); and (2) *Developmental and Health Screenings*, which characterizes the use of screening tools to identify any health or developmental issues.

Child Observation

Though providers generally agreed the Child Observation element was important, opinions on the DRDP and DRDP Tech were varied.

Providers recognized the importance of observing children to monitor their learning and development—the underlying goal of the Child Observation element. However, strong opinions on the DRDP and especially the DRDP Tech overshadowed the conversations. A few providers mentioned the benefits of using the DRDP and/or the DRDP Tech. "The DRDP was very helpful for me," one provider emphasized. "Using all those tools really helped organize my situation a little bit more. Everything else was a little less chaotic, and the DRDP wasn't a chore. It was more like I could just see it evolve in front of me, and the notes were very easy." Another provider said that of the assessment tools, the DRDP was most helpful, adding that she was really glad to see that the DRDP Tech was "going to be free for state-funded programs." This same provider noted, "What it takes to do the DRDPs…even though it's good to know where everybody is, it's also hard and distracting for the teachers to always be planning and doing it. So

I think having the DRDP on the computer will make it a little simpler for them, and it's going to improve the quality. I think it's going to simplify it and that it's been a thorn in most people's sides for a long time."

Despite some positive feedback about the DRDP tool, some providers expressed significant challenges related to DRDP Tech. For example, several providers noted that the cost of the technology needed to implement the DRDP Tech would deter them from trying to reach a higher tier. As one center director explained, "We chose not to spend money for the DRDP Tech, so obviously that rating can't go up until we make that purchase. But we've decided that we're okay with that."

Most parents appreciated the value of the Child Observation element, although many did not fully grasp its importance until after they had enrolled their children in care; for others, this was not a high-priority indicator of quality.

Parents in 15 of the 17 focus groups indicated their child had been assessed and the results had been shared with them. One parent explained how the curriculum and assessment were linked: "They'll show you that they're teaching the triangles, the shapes, they have a little folder with all the stuff they work on, and you can see the progress of what they're working on. You can see the difference from when they first started." Another parent, pleased to see the results, stated, "I'm very impressed with what my child has learned in a short period of time…. For me, what he has learned has surpassed what I expected out of childcare."

Although parents overwhelmingly agreed the Child Observation element was important, in 10 Consortia, they reported that they only fully understood the importance after they had enrolled their child in care. As one parent explained, "When you start at the beginning, you aren't thinking about curriculum and development…. I was looking basically for safety in a loving, happy classroom…. I didn't see [the importance of measuring children's learning] before."

A few parents were unclear on how the assessments worked or did not feel that they were an important aspect of their early learning program. One parent explained, "I was looking for nurturing and just simple things, not thinking they were going to 'test' them." Similarly, some parents, unfamiliar with how the assessments were used, felt "tests" could not be accurate or appropriate for young children. This may indicate that the element should be more clearly articulated to convey that educators are measuring children's progress and adapting curriculum and learning opportunities to support children's development.

Developmental and Health Screenings

Although several providers found the developmental and health screenings to be beneficial, others reported facing challenges conducting the screenings and/or reviewing assessments with parents.

Several providers had positive feedback about the benefits of using developmental and health screenings. For example, one family child care provider noted that the Ages and Stages Questionnaire (ASQ) has improved her communications with parents, adding that having the tool to go over during parent conferences has been really valuable to both her and the parents she serves.

In contrast, two early learning staff members reported that the annual physicals needed for the Health Screening Form were a challenge, particularly for families without health insurance. However, one provider noted that "it was challenging to have parents understand why…it was so important to take their child to the physician on a yearly basis to have their health checked. So that was a teaching part for our families." In another Consortium, a center director noted that the Health Screening Form had been one of the challenges their site had faced, but that the experience had improved with each passing year and that there was an increase in alignment among systems. This center director explained, "Each year it gets better. Each year it's a more mainstream system in which we can connect with either the school district around us or the health screenings. It gets a little easier…. I feel like we're working more together than in the past. I see improvement. Not completely there, but getting there."

Several providers discussed the difficulties of implementing the child screenings and reviewing the assessments with parents. One center director interviewed explained: "The ASQ was a little challenging for us but it provoked us into learning it and using it. It's a little challenging for parents too…helping parents understand how to use the tool and how to have the conversation."

The Developmental and Health Screening element was not a priority element for most parents, although some parents, especially those with children with developmental delays, recognized its importance.

In 15 of the 17 parent focus groups, the parents were aware that their children were screened and had received results. As noted above, use of ASQs was not necessarily the primary factor in choosing care, although several parents noted that they better understood how critical it was during or after enrollment. For example, one parent noted, "Maybe initially it wasn't at the top of my list of what was important because I was looking at price and location, but when I went there and actually talked to [the child care provider] and found out this was really important to her, and having a two-year-old that was also having developmental delays…that really sealed the deal for me. Yes, this is where my kid should go, because she's going to help and she's going to refer her to other services she may need." Parents of children with delays especially appreciated screenings and tracking children's progress. One parent reported, "That's how I was able to notice my daughter's [speech] problem. I would have never noticed because it's my daughter and she just talks funny to me. They're working with her; I don't have to do anything. They say 'we will do this'." Another parent cautioned against over-assessment and the risk of over-diagnosing attention deficit hyperactivity disorder (ADHD), however: "I think it's good but you don't want it to be too much. They're really young."

Core II: Teachers and Teaching

During interviews and focus groups, providers and parents also provided feedback on *Core II: Teachers and Teaching,* the second domain of the Hybrid Rating Matrix. *Core II* includes two elements: (1) *Minimum Qualifications for Lead Teacher/Family Child Care Home (FCCH)*, which focuses on teacher qualifications, including unit and degree attainment and other professional development opportunities; and (2) *Effective Teacher-Child Interactions: CLASS Assessments*, which focuses on teachers' familiarity with the Classroom Assessment Scoring System (CLASS) at the lower end of the QRIS rating scale and the teacher-child interactions as score measured by CLASS assessments at the higher end.

Minimum Qualifications for Lead Teacher/Family Child Care Home

Although the majority of the providers agreed with the requirements for staff's professional development hours, the most commonly cited challenge related to the rating matrix concerned achieving the college degree requirements.

Among the providers interviewed, the most commonly cited challenge related to the Hybrid Rating Matrix was attainment of the degree requirements. Ten providers noted that the degree requirement was difficult; of these providers, five were classroom teachers, three were family child care owners, and two were center directors. "It's hard to find staff that has their Bachelor's [degree]," said a center director. "And sometimes they don't have their A.A. degree."

On the other hand, the majority of providers reported that the 21 hours of professional development was an appropriate and effective amount of professional development to improve quality, although some pointed out that the quality of the training was more important than the quantity of hours. In contrast, several providers felt that the required hours were difficult to attain for various reasons, including scheduling conflicts and the fact that providers are being asked to attend training events on nights and weekends, which may conflict with work schedules and personal commitments.

Many parents were familiar with their teachers' qualifications and considered them when selecting a program for their child; however, although some felt strongly that teachers should have degrees, most parents did not consider teacher qualifications as the only—or even the best—indicator of quality.

In the majority of the parent focus groups, at least one parent reported being familiar with the training and qualifications of the teachers in their child's classroom. Parents had found out about teachers' qualifications in various ways. In some sites, the information was included in parent letters, posted at the site, or available online. In other sites, parents had learned about the staff's qualifications directly from the staff themselves or the center director. In contrast, in 10 of the 17 parent focus groups, at least one parent reported not knowing the teachers' training and qualifications. And some parents did not know how to find out, as one parent described: "When we were searching, I didn't even know how to officially look to see what quality training they'd taken…. [I] didn't know what to look for."

All of the parent focus groups included parents who said that they had considered teacher qualifications when choosing child care. A few felt very strongly that their child's teacher should have a college education. For example, "I know there are schools that will hire you with 12 ECE [early childhood education] units or less, some even 6," reported one parent. "That's two classes. I'm sorry, but you're not going to watch my kid after just two classes. [That is] one of the major things for me."

Generally, parents who indicated the teacher qualifications were important stated that they wanted a teacher who knew what he or she was doing and who knew how to teach their particular child. Several focus group participants noted that they had a child with special needs who benefitted from a trained teacher. "Knowing she had those credentials, we had more confidence knowing that [our child] would be learning what he needed to learn," explained one

parent whose child had a teacher with a Bachelor's degree and specialization in special education.

Eight focus groups included parents who said that they did not consider teacher qualifications—at least not as they are described in the Hybrid Rating Matrix—when choosing an early learning and care program for their child. Among those focus group participants who indicated that teacher qualifications were less important when seeking care, parents expressed concerns that a college degree is not the only measure of quality and that the teachers' experience and connection to the children should be considered first. One parent explained, "I don't think that I looked for that actually. It's a good thing to have, but when I was searching for a school, I wanted someone who could communicate with my children, not that had all the certificates."

Effective Teacher-Child Interactions: CLASS Assessments

Although most providers reported benefits of the CLASS element, parents' understanding of and feedback on the CLASS tool were mixed.

Providers' feedback on the CLASS element and tool was primarily positive. For example, providers frequently noted the influence of the CLASS on their teaching practices. As one center director reported, "The most useful [tool] is probably the CLASS tool…. I think it really provided a focus on the interactions, on what's really happening in the classroom." A family child care provider said that she liked the CLASS because it is "precise" and offers "structure and guidance."

In 8 of the 17 parent focus groups, parents were aware of teacher observations. Parents' awareness of who was conducting these teacher observations varied. For example, although some parents were aware that they were being conducted by external observers, many parents were not sure who was conducting the observations.

Parents generally valued the role of a good teacher in the classroom but did not necessarily believe that the classroom observational assessments aligned with their priorities. For example, one parent explained, "For me, originally, it was more somebody I could trust more than what they were able to teach…. Eventually, as the child gets older, you want more and more of that, but at the infant stage, it was more about being able to go to work and not worrying about my child for 10 hours."

Core III: Program and Environment

During interviews and focus groups, providers and parents provided feedback on *Core III: Program and Environment*, the third domain of the Hybrid Rating Matrix. *Core III* includes three elements. First, *Ratios and Group Size* addresses the total number of children in the classroom as well as the number of children per teacher; this element is not scored for FCCHs. Second, *Program Environment Rating Scale(s)* focuses on the environment of the early learning and care program, as measured by the use of the Environment Rating Scales (ERS) tool at the lower end of the QRIS rating scale and the ERS score at the higher end of the scale. Third, *Director Qualifications* includes both unit or degree attainment and other professional development experiences for center directors. Family child care providers are not included in this element.

Ratios and Group Size

Some providers reported that attaining the needed staff ratios was a major barrier to moving up the tiers, and although some parents highly valued a more protective ratio, for many, this was not a priority.

Several center directors spoke about the difficulty of attaining the staffing ratios for the higher tiers given their current funding. These center directors often worked in a state-funded preschool and/or at a school district, and they met their ratios using parent volunteers. The QRIS requires staff, not parent volunteers, to fulfill the ratio. As one center director explained, “I can’t go down to 2 to 20 in the program…because our funding doesn’t allow for that.” Another director noted, “We can’t hire another instructional assistant. We use parent volunteers. We can’t hire using money that comes and goes…. It has to be sustainable.”

Parents varied in their awareness of the staff–child ratio in their child’s program. In 14 of the 17 focus groups, parents indicated that they were familiar with the teacher–child ratio, but in 7 of the focus groups, some or all of the parents indicated that they did not know their program’s teacher-child ratio.

In 11 of the 17 parent focus groups, parents discussed their perspectives on the importance of the teacher–child ratio. Although some parents understood the importance of the teacher–child ratio, it seemed to be a deciding factor for only a few of these parents. As one parent explained, “When I went to enroll my child, I needed to know how many teachers in each classroom and what the maximum limit of kids was. Because I knew my child was going to need more support.” Another parent, pleased with the ratio at her program, reported, “I know it will impact my child tremendously because they get a lot of attention here.”

However, the ratio was less important for some parents. As another parent shared, “For the majority of us, what matters to us is to find someone to take care of our child so we can go to work. Later, when there is time, you start to figure out all the things about the center.” Another parent echoed this limited interest in (and understanding of the importance of) staff–child ratios: "I don’t know, I just know that they are teachers. As long as I see a bunch of kids and some teachers, that’s okay with me."

Program Environment Rating Scale(s)

Provider views of the ERS element were mixed, with some seeing the benefits of a focus on the environment, and others focusing mainly on challenges concerning the implementation of the assessment; parents were mostly unfamiliar with the ERS element.

Feedback about the ERS varied among providers. In some cases, feedback was positive. For example, one center director shared that prior to use of the ERS, the teachers at the site “knew the basics [about classroom environments], but once they started to read ECERS [Early Childhood Environment Rating Scale], they started reading more, getting more information. Now they’re going to their classrooms knowing more about how things need to be set up. It really helps them.” Another center director noted the benefits of using the CLASS and ERS in tandem, adding, “Some of the tools are helpful, but seeing it all in one place.... There are research-based ways to measure our effectiveness. Before, we would get an ECERS score and a

CLASS score, but when we put them all in one place, it was easier to see where we could improve a bit."

Other providers expressed challenges related to implementation of the ERS. As one center director explained, "Some of the areas are out of our control, which makes it a little difficult. And then, as the years go by, there are different people that come out and suddenly something that was fine last year, they'll come out the next year and say 'No, that's not okay.' So it's just consistency." Another center director noted that although the site has been using the ERS and CLASS for some time, since Race to the Top-Early Learning Challenge (or RTT-ELC) "there is more work in monitoring it to see if it works. It is understandable, but it adds a lot of extra time. [The administrative agency] has a lot of requirements. I guess we are providing the same information in a couple of different ways now. It is a duplication of reporting information."

In 11 of the 17 parent focus groups, parents were not familiar with the ERS element or the fact that the program environment was evaluated. And in five other groups, some of the parents were familiar with the ratings and others were not. The few parents who knew about ERS tools often worked or volunteered in the program. As one parent explained, "I knew about ECERS because they said we need a parent from each class to go to a meeting and I went to the meeting and they talked about ECERS. But it was only for the parents who volunteered."

Director Qualifications

Director qualifications were generally viewed as important, although providers noted that it was difficult to score high on the element, and parents were generally unaware of their director's qualifications and unsure of how to evaluate them anyway.

Providers noted that director qualifications were an important consideration for the rating matrix. Some reported that it was difficult for many center directors to attain the highest points in the Hybrid Rating Matrix for this element, however. One center director observed, "I think the top-tier education levels are hard but you can't take those off. I have my master's degree, but I know a lot of [center directors] don't."

Parents agreed that having a well-qualified center director is important. For example, one parent highlighted the point, "A qualified director—it all starts with a director. If she's not doing her job, the teachers aren't doing their job." But in 12 of the 17 parent focus groups, parents did not know the program director's credentials. "We just trust the school district hires the right people," said one parent. "That they went to school. That they know what they're doing. That they have education. That they have credentials and they're good for these kids at this age." Some parents indicated that they did not know what to look for in terms of director qualifications. For example, one parent said, "They say 'I've been here for 20 years or I've been here for 15 years.' That's all, but we don't even ask them, we don't even know what to ask."

Other Elements That Influence Perceptions of Quality and Parent Choices

Additional Quality Considerations

Providers and parents also discussed additional quality factors—beyond those delineated in the rating matrix—that impacted their perceptions of child care. Parents' observations of these characteristics were often based on what happened after their child began care rather than as part of their selection process.

Providers and parents both highlighted the importance of family engagement as an important indicator of program quality.

Several of the providers had a broader perspective on quality elements, having participated in earlier quality initiatives and/or accreditation efforts, and commented on the need to include a measure for family engagement. "There is nothing about the parents' role except on the ASQ," one teacher explained. "It isn't measured, but I think it is important for quality." A center director explained, "I think having family engagement plays a big part in the success of the preschool. If the families are invested, that's obviously going to have an effect on the child and the center. If you were involved in some of the previous parts of it, like Preschool for All, you might have seen it, because family engagement is one of the things they have measured."

Parents also referenced the extent to which program staff engaged them in the program and their children's learning. This was discussed in 11 of the 17 focus groups. One parent described teachers' effective dissemination of information, "What I like a lot is that the teachers keep parents well-informed. They post announcements frequently about everything that's happening at the school, the gatherings at the school, how the children are behaving, and what the child needs help with—lots of information. In both languages." Another parent emphasized the importance of communication, "Feedback is a two-way communication; it's helpful to have the feedback about what happened to your child this day but also for you to give feedback to the teachers." Similarly, another parent added, "So I would say collaboration between parent and teachers, and having the ability to think outside the box and use different resources and interventions to meet all of the children where they are."

Providers and parents suggested alternative perspectives on the Teacher Qualifications element to make it more appropriate and/or attainable.

Four providers discussed alternative ways to look at teacher qualifications. Several noted the need for a way to better assess the warmth or nurturing that early learning staff provide to children, even if the education requirements are not met. One FCCH provider explained, "I know there are some child cares where the teachers aren't educated but they're nurturing in other ways…. I think the scale doesn't measure how nurturing a teacher is—that should be captured in the rating."

Like many of the providers interviewed, parents also mentioned the importance of having a kind and nurturing teacher. This topic came up in 15 of the 17 parent focus groups. "If a teacher is treating my child like they would their own, then that's quality." Although the Emotional

Support domain of the CLASS captures the extent to which teachers have positive, supportive, and responsive relationships with the children, providers and parents described a more elusive characteristic capturing a teacher's kindness or likeability—something that could not easily be added to an objective rating scale, but is important to families nonetheless.

Other providers pointed to the limitations of the CLASS for adequately capturing all of the important aspects of classroom interactions and suggested that this element could be expanded. One director explained, "With the CLASS piece, while I like it, there are some pieces where you could still score well and yet not necessarily have the best practices.... A teacher could be talking and presenting things in not necessarily a developmentally appropriate way…. I think more attention of what actually happens in the classroom [is needed]…."

Many parents pointed to child outcomes—their children's learning and happiness—as important indicators of a quality program.

Parents in 16 of the 17 Consortia mentioned that their child's positive experiences at the site were the most important indicators of quality. Parents in all of these focus groups specifically mentioned that their children's developing school readiness skills were an indication of a quality program. One parent noted, "I love when my kid comes home with all his artwork and is practicing his writing and the art and all the songs he knows." Learning social and behavioral skills also was important to parents in 6 of the 17 focus groups. One parent described, "The kids have learned how to sit still, how to follow directions, how to get in a line." This is important, the parent continued, "Because if they can't sit still to take in the information then they're not going to learn." Seeing their children happy in the program also was an indication to parents that they had selected a quality program. As one parent put it, "If my child likes it, is comfortable when I drop him off, and doesn't want to leave when I pick him up, then it's a quality program."

Although parents often did not know how to evaluate the curriculum, they valued the importance of a good curriculum.

The use of a quality curriculum was mentioned by parents in 11 of the parent focus groups. Parents often did not know how to evaluate the curriculum, but they liked what they saw. "There is a structure, and a good curriculum and a developmental path that they're using," described on parent. "For me, it's developmentally appropriate practices," explained another parent. "Knowing that what it is they are putting before the child is not above the child's head but something obtainable, something they can do."

Providers indicated their desire for flexibility in QRIS standards, considering the variation across sites.

Some providers interviewed noted that they would like to see a little more flexibility in the standards to accommodate the realities of the resources and sites. This was especially true among FCCHs. Most of the FCCH providers interviewed expressed frustration with the current QRIS standards. As one clearly stated, "You can't apply the same rules to centers and family child care homes, they are two very different animals." Another suggested that instead of a one-size-fits-all approach, there should be "more leeway for family child cares and acknowledgement that there are few resources and very diverse needs—different ages, specials needs, etc."

Center directors also indicated there were standards that did not fit their centers. Two centers noted that they operate on a school site and/or are part of a school district, and the rules that regulate their sites do not always match the requirements of the QRIS. These issues, which often involved environmental factors, were viewed as being outside of their control and prevented them from moving up the tiers.

Others suggested that there might be an alternative pathway to earning points on the Minimum Qualifications element for those individuals who are unable to achieve a Bachelor's degree. Interview respondents pointed out that there may be providers who are better qualified to provide high-quality care, but they are penalized because they don't have the Bachelor's degree, and, given their circumstances, getting the degree is not feasible. One FCCH provider suggested that exceptions to the degree requirements could be made if CLASS and ERS scores were high.

Other Factors That Contribute to Parents' Choices

When choosing an early learning and care program, parents were most likely to indicate that they relied on factors outside of quality indicators such as their own comfort with the site, the convenience of the site to their home or work, and the cost of care.

Parents described comfort with the site and staff as an important factor in selecting a program for their child.

In 15 focus groups, parents spoke of choosing an early learning and care setting based on their own comfort with the site and staff. One parent explained, "It was more important that my child was safe and was in a good environment and I could trust the people that were watching my children. For me, you have to go with your gut instinct, and my gut said that I loved it." This may be related to other elements such as the ERS element, but it was described more as a feeling and less as something that could be quantified. Numerous parents noted that they looked for "somebody I could trust." Parents reported that meeting face to face and talking with the staff about the program gave them the reassurance they needed to make their selection.

Parents identified convenience of the site to their home or work and program schedule as well as cost of the care as key factors in their decisions about selecting a program for their child.

Parents in 14 of the focus groups mentioned convenience as one of the factors that guided their care choice. Some spoke about wanting care for certain hours and others talked about the proximity to home or work. One parent described her selection decision, "I was right around the corner and I didn't drive so it was walkable."

Cost was noted in 11 of the 17 focus groups as an important factor in choosing a child care arrangement. For some the key selling point of a site was that it was free; for others, they selected a site where they could afford the fees. "I was mainly looking for cost," explained one parent. "[I] didn't want it to be too expensive because I have two children, so I needed it to be either income-based or something that would fit into my budget."

Publicizing Ratings

Providers and parents also had views about the dissemination of ratings to the public. Parents expressed their interest in this information and also described the sources of information they actually used in making their decisions—in the absence of the ratings.

Parents currently rely on a variety of sources to gather information to inform their choices about early learning programs, including recommendations from family and friends, and online resources.

The most common source of information to guide care choices reported by parents was a recommendation from another parent, a family member, neighbor, or coworker. Parents in 16 out of 17 focus groups reported they rely on personal recommendations. Parents rely on people they know and can trust to tell them what programs are of good quality: "That personal reference made me feel okay," said one parent about her reliance on word-of-mouth recommendations.

The second most common source of information was online or print media, reported in 11 out of 17 focus groups. For example, some parents reviewed local parenting resource magazines (e.g., *Bay Area Parent*) for listings of programs. But many relied on the Internet, through Google searches, Yelp reviews, and advertisements on Craigslist or Care.com. Parents noted that they appreciated the reviews—reactions from real parents who had positive (or negative) experiences about a program and shared them online. Less common responses included the Resource and Referral network or their local First 5.

Parents were interested in learning the ratings, and, for the most part, wanted detailed rating information.

Parents in all focus groups indicated that they would like to have access to the rating information, and more information is better than less information. In all 17 parent focus groups, at least one parent noted that the individual element scores were more useful than an overall score. As one parent explained, the individual scores "allow you to assess your own hierarchy of what's important to you and whether [the program is] scoring well in the areas that are important to you." Another parent added, "I'm sure each family has different priorities…. It's good to have a different score for each one, because [if] you can't have [the highest] quality in all of them…at least you can pick which one is more important."

Many felt that one overall score would obscure the important information. "One overall rating is like what they use in the school district," said one parent. "That's not really telling you anything." Similarly, another parent indicated that "Every child is different, so you need to pick according to what's going on with your child…. If you had an overall score, how would you know?"

In contrast, a few parents admitted they would be content with just one overall score. "I think a general, an overall score is better," expressed one parent. "It's all included in there. If you start to analyze each item, you'll never be happy."

Some providers saw the benefits to informing parents about the ratings, while others questioned the accuracy of the ratings.

Several providers expressed various benefits from publicizing ratings. Two of the family child care providers thought the rating would give them legitimacy and help improve parents' understanding of their child care program. In addition, two providers noted the need to share the successes of centers in the highest tiers, both to motivate sites to improve and to inform parents about the need for quality. As one provider argued, "I know that the group has been very reluctant to post the scores of programs because they don't want to discourage the programs in the lower ranges…. The message is that you are in the program, and that means you are committed to quality. But there hasn't been recognition at the top. There isn't an incentive for the programs with a '3' or a '4' to move up. We can't say we are at [Tier 4]. This is great for the lower tiers; it takes off the pressure on the lower-rated programs to participate, but it doesn't give much incentive to go higher. To really promote quality in the field and the in the county, the parents should really know that information."

In contrast, some providers expressed frustration with the ratings and felt they didn't accurately reflect the quality of their program.

Although ratings are, for the most part, not publicly available, many of the providers interviewed reported that they share information with parents about their involvement with the QRIS—the process of being rated and/or their participation in quality improvement (QI) activities.

More than half of the respondents indicated they did share information about their participation in the QRIS with parents, that parents are interested in the information, and that the program benefits from informed parents. One provider reported that it is helpful "anytime you can educate parents about what a quality program even is." A FCCH provider noted that providing some information gave parents a better understanding of the quality of the program their child was attending: "If I put it out there in my handbook and in my other paperwork that this is what we're providing here, which is what I've done, I think it lets them know that the program is improving, I am improving, so things are improving for their child and their care here." Some providers felt the parents really appreciated the information, while others were concerned that the parents were simply not interested in their QI efforts. "I think maybe it matters to some parents," explained one FCCH provider, "but to some I don't think they really care." Some providers were concerned that it was just too complicated for parents. Others only provided the information if parents ask, and they haven't asked yet.

A few providers agreed that it would be important to share information about their QI efforts but had not yet done so. One center director explained, "We haven't gone to that step yet. It's not that we don't want to; it just hasn't become part of our process yet…. We're just trying to…use it internally and the value that it brings. It hasn't become the priority yet to work on it with families."

Summary

During the interviews and focus groups, providers and parents generally agreed that the Hybrid Rating Matrix included the important aspects of quality and generally measured the appropriate elements of quality. Although they did not think that any of the elements should be eliminated, some providers and/or parents took issue with particular requirements in some elements. For example, the most common challenge cited by providers was the college degree requirements for teaching staff.

Knowledge, understanding, and prioritization of the elements of quality early learning and care programs varied among parents. For example, many parents were not aware of or did not understand the importance of particular elements of quality care (such as child assessment) until *after* they had enrolled their children in the program. Prior to enrollment, many parents prioritized other factors than the elements included in the rating matrix, such as cost, proximity of care to the parents' home or workplace, or their degree of comfort with the staff and program as a whole.

Providers and parents had different perspectives on making the ratings publicly available—not surprising given their different roles. Several providers expressed various benefits from publicizing ratings, including the fact that it could help legitimize their program and improve parents' understanding of their early learning and care program, but others expressed frustration with the ratings and felt they did not accurately reflect the quality of their program. In contrast, parents in all of the focus groups indicated that they would like to have access to the rating information, and they generally noted that individual scores would be more useful than an overall score.

Finally, providers and parents also discussed additional factors—beyond those delineated in the rating matrix—that captured the quality of an early learning and care setting, and that could be considered for potential inclusion in the rating matrix. Both providers and parents emphasized the importance of family engagement as an indicator of quality. Providers and parents also suggested alternative perspectives on the Teacher Qualifications element to make it more appropriate and/or attainable, such as whether the program scores highly on the CLASS and/or ECERS, which could be an indicator used in place of requiring teachers to obtain a college degree. Finally, parents in nearly all of the focus groups mentioned their child's positive experience in the program and their increasing school readiness as important factors that indicate a quality program from their perspective.

Chapter 5. Staff-Level Quality Improvement Supports

Key Findings

This chapter provides descriptive information about early learning teaching staff participation in quality improvement (QI) activities, drawing on surveys of 406 early learning staff from the 142 centers and family child care homes (FCCHs) participating in the outcomes study.

- Center-based staff surveyed reported substantial engagement in QI activities, and many teachers reported consistent QI activity participation over the school year.
- The majority of center staff (80 percent) reported receiving coaching and mentoring supports, and staff reported an average frequency of 18.3 coaching interactions over 10 months.
- Approximately three quarters (73 percent) of lead and assistant teachers in centers reported that they participated in noncredit workshops or training, on average, for about 28 hours over the year.
- Almost all center staff reported spending some coaching and training time on language development and literacy, math/cognitive development, and social and emotional development.
- More than half (57 percent) of center staff reported that they participated in formal peer support activities such as learning communities, peer support networks, or reciprocal peer coaching, with an average of 22.8 total hours of peer support time.
- Just over one quarter (27 percent) of center staff reported that they had participated in credit-bearing coursework at a college or university mostly focused on early childhood education.

The descriptive analyses are not representative of all programs in California, or of all programs participating in the California QRIS. They represent a self-selected sample of fairly well educated staff in programs participating in the California QRIS.

This chapter provides descriptive information about early learning teaching staff participation in quality improvement (QI) activities and addresses the following research question:

- RQ 9. What are early learning staff's experiences with quality improvement activities?

The findings presented in this chapter are descriptive in nature. The data on which they are based are also included in further analyses discussed in chapter 7. The data were collected through an online or paper survey of 406 early learning staff from the 142 centers and family child care homes (FCCHs) participating in the outcomes study. They represent a self-selected sample of

respondents participating in Race to the Top–Early Learning Challenge (RTT-ELC) grant programs within the study's focal Consortia. The picture we draw here may not reflect all ECE staff and indicate underlying variation in who chose to participate in the Quality Rating and Improvement System (QRIS) and respond to this survey. As a descriptive study, we are interested in documenting patterns among QRIS staff, although it is important to note we are not able to explain the reasons for the differences we observe or analyze variations in the quality of the activities. The survey results are supplemented with information gathered from interviews with a small number of staff to add texture to our understanding of early learning staff's experiences with QI activities.

Analysis Approaches

- Conduct a descriptive analysis of early learning staff survey responses.
- Report question response percentages and mean values as appropriate.

Data and Sample

- In total 406 lead and assistant teaching staff representing 234 classrooms across the 142 sites participating in the outcomes study were surveyed.
- Respondents include 306 lead and assistant staff in centers and family child care homes across the 11 focal Consortia, a response rate of 75 percent.
- A total of 81 percent of respondents are from centers and 9 percent are from family child care homes.
- About two thirds (66 percent) of respondents are Hispanic or Latino, and another 12 percent each are non-Hispanic White or Asian. Almost two thirds (63 percent) of respondents had attained an Associate's degree or higher; 40 percent report a Bachelor's degree or higher (see appendix exhibit 5B.2).
- Respondents represent a fairly new workforce, with 65 percent reporting five or fewer years' experience working with young children.
- English is the primary language for about half (51 percent) of respondents, and 40 percent reported Spanish as their primary language.
- The age range of respondents is fairly evenly distributed from age 20 years to over 60 years.
- By comparison with data from a 2004 California statewide early learning workforce sample, the QRIS study sample has higher reported rates of staff who are Hispanic or Latino, who are in older age brackets, and who have a Bachelor's degree or higher (see appendix exhibit 5B.3).

Staff Sample Demographics

Of the 406 teachers in our final survey sample, we received complete survey responses from 306 (a 75 percent response rate): 189 lead teachers[9] and 117 assistant teachers. The majority of the 306 responding teachers worked in centers (81 percent, or 279 staff compared with 27 staff working in FCCHs) and with preschool-age children (80 percent or 244 staff). They represented all 11 focal Consortia (see appendix exhibit 5B.1).

Respondents from FCCHs represent only 9 percent (27 staff total, 15 leads, and 12 assistants) of our total survey sample. Their distribution across Consortia was uneven: 70 percent of all FCCH respondents work within two of the 11 Consortia, and four Consortia do not have any FCCH staff represented in the survey sample. Given the small FCCH sample and the lack of representation across all Consortia, we focus this chapter on center staff, breaking out results for lead and assistant teachers. We report at the end of each results section how the FCCH respondents answered, but

[9] For the purpose of this chapter, the term *lead teachers* includes both lead and coteachers. *FCCH staff* includes both leads and assistants. For a detailed description of survey methods and sample, see appendix 1A.

we do not break out FCCH staff by leads or assistants. Given the very small sample size, we note that these FCCH survey results should be interpreted with caution.

Participation in Quality Improvement Supports

To begin the survey, we asked staff to indicate whether or not they had received four types of QI support from June 2014 through March 2015 to improve their practice or program quality (see sidebar for definitions of each QI type). If a respondent answered yes for any type of QI support, we then asked a set of questions specific to that QI activity.

Of the four QI types, more center staff (80 percent in total) reported receiving coaching or mentoring supports than any other type of QI support (exhibit 5.1). This held true for both lead teachers and assistant teachers (82 percent and 77 percent, respectively). The next most often reported QI type was noncredit workshops or training, with almost three quarters (73 percent) of staff reporting participation (76 percent of lead and 66 percent of assistant teachers).[10] Lead and assistant teachers reported participating in peer support activities at the same rate (57 percent). Center staff reported participating in credit-bearing courses the least (27 percent in total; 24 percent among lead teachers and 32 percent among assistants).

QI Support Definitions

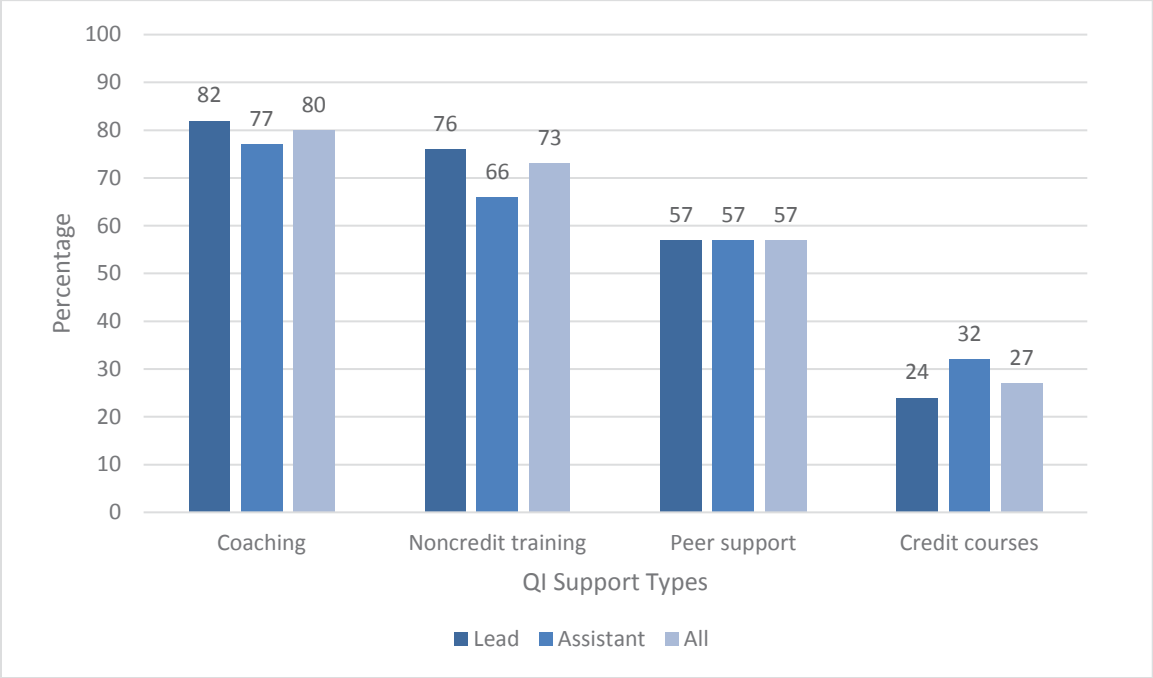
Respondents were asked about these types of QI support, defined in the survey as follows:

- Coaching or mentoring supports: Supports for individualized professional development, usually one-on-one or as part of a classroom team, provided by a coach, mentor, or advisor to help improve staff practice or to promote quality improvement more generally.
- Noncredit courses, seminars, workshops, or training programs: A training activity that may be one-time or part of a series (including courses that provide continuing education units but not including courses taken for formal college credit through a college or university). This QI type is identified as "workshops or training" throughout the remainder of the chapter.
- Peer support activities: Formal arrangements such as learning communities, peer support networks, or reciprocal peer coaching to discuss shared experiences and exchange ideas, information, and strategies for professional development or for program improvement more generally. Informal or occasional discussions with colleagues were not to be included.
- Credit-bearing college or university courses: Course(s) completed for unit credit at a two- or four-year college or university.

Among the FCCH respondents, staff reported participating in coaching and noncredit workshops or training most often (85 percent each) among the four types of QI support, as shown in appendix exhibit 5B.5. Almost half of FCCH staff reported participating in credit-bearing courses and peer support activities (48 percent each).

[10] The findings discussed in this chapter are descriptive only and do not imply statistical significance. Any differences cited here or in subsequent sections reference comparisons within the survey sample only.

Exhibit 5.1. Percentage of Center Staff Who Received Quality Improvement Supports (June 2014–March 2015)



SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.4.

NOTE: The sample includes a total of 279 respondents, with 174 lead teachers and 105 assistant teachers. Lead teachers include lead teachers and coteachers.

The following sections of this chapter describe staff responses to the more detailed questions posed when respondents indicated they had received a particular type of QI. These included questions about dosage and content areas, as well as questions about financial incentives, perceived helpfulness of the QI type, and factors influencing how staff members decide to participate in QI activities. Note that appendix 5A provides the survey questionnaire, and appendix 5B includes complete exhibits of the data referenced in each of these sections.

Coaching or Mentoring Supports

Dosage

Most center staff reported receiving coaching and mentoring supports, although reported monthly dosage was not necessarily consistent across months.

Most center staff reported receiving coaching or mentoring supports at some point between June 2014 and March 2015, although receipt was not necessarily consistent over all months. Eight out of 10 (80 percent) lead and assistant teachers reported that they had received at least some coaching or mentoring during this period. This is more than the percentage of staff (70 percent) who reported in the same survey receiving coaching or mentoring the *previous* year (June 2013 through May 2014). When considering the percentage of staff who reported receiving coaching

or mentoring at any point during the combined two-year period (June 2013 through March 2015), the total rises slightly (84 percent). See appendix exhibits 5B.4 and 5B.6 for details.

Of staff who received coaching or mentoring from June 2014 to March 2015, lead teachers reported receiving a total average of 22.9 hours per person across the 10 months and assistant teachers reported receiving 21.0 hours (exhibit 5.2) Lead teachers reported an average *monthly* number of 2.3 hours of coaching or mentoring per person and assistant teachers reported 2.1 monthly hours. Coaching or mentoring time was reported to be lower in the summer months (June, July, and August), and lead teachers reported the highest average monthly hours (3.2) in the month of September. These data can be found in appendix exhibit 5B.7. When examining consistency of this QI support over time, we found that half (49.5 percent) of total center staff reported at least some coaching or mentoring time during every month from September 2014 through March 2015 (see appendix exhibit 5B.9).

Exhibit 5.2. Coaching Hours and Frequency per Person by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
Hours			
Total average hours from June 2014–March 2015	22.9	21.0	22.2
Average hours per month	2.3	2.1	2.2
Frequency			
Total average frequency from June 2014–March 2015	18.6	17.7	18.3
Average frequency per month	1.9	1.8	1.8
Number of respondents	142	81	223

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibits 5B.7 and 5B.8.

We also asked staff about how many times they received coaching or mentoring in each month. From June 2014 through March 2015, staff reported an average frequency of coaching of 1.8 times per month and 18.3 times across the 10 months.[11] Within each month during the school year (September through March), 15 percent to 20 percent of lead teachers reported receiving coaching support three or more times. Not surprisingly, these percentages are lower for the summer months, when the majority of center staff reported that they did not receive coaching or mentoring at all. Less than 6 percent of all staff reported receiving coaching or mentoring support five or more times for any given month. See appendix exhibit 5B.8 for reported frequency by month.

[11] Frequency of coaching was calculated by averaging assigned midpoint values to categories asked in the question. Appendix exhibit 5B.8 shows percentages of respondents for each category. Categories of frequency (and corresponding midpoint values) included: "Not at all" = 0, "1 or 2 times" = 1.5, "3 or 4 times"= 3.5, and "5 or more times" = 5.

Content Areas

Almost all center staff reported spending some coaching or mentoring time on language development and literacy, math/cognitive development, and social and emotional development. The top five *additional* content areas reported by staff align with elements of the CLASS or ERS assessments.

When asked what percentage of coaching or mentoring hours was spent on three core content areas—language development/literacy, math/cognitive development, and social and emotional development— almost all center staff reported spending some time on all three content areas. A total of 98 percent of staff reported receiving any coaching or mentoring on social and emotional development, 96 percent on language development/literacy, and 94 percent on math/cognitive development. The largest percentage of coaching or mentoring hours was spent on social and emotional development. Nearly half (47 percent) reported that half or more of their coaching time was spent on social and emotional development, while only 21 percent reported that half or more of coaching time was spent on math. See appendix exhibit 5B.10 for details.

We also asked staff about 20 *other* potential content areas addressed during the coaching or mentoring support they had received (exhibit 5.3). For lead and assistant staff alike, the top five content areas reported included teacher-child interactions, materials and learning environment, and child behavior management—all of which align with elements of the Classroom Assessment Scoring System (CLASS) or Environment Rating Scale (ERS) assessments. For lead teachers, understanding and improving scores on CLASS and ERS also were among the top five, and child assessment and developmental screening and health and safety rounded out the top five for assistant teachers. The content areas *least* reported by all staff were business practices, accreditation, and licensing issues. Appendix exhibit 5B.11 provides the response rates for the full list of content areas included on the survey.

Exhibit 5.3. Top Five Most and Least Reported Coaching or Mentoring Content Areas by Staff Type: Centers (June 2014–March 2015)

Lead		Assistant	
Most Reported			
Content area	%	Content area	%
Understand/improve scores on CLASS	83.8	Teacher-child interactions	78.8
Teacher-child interactions	83.1	Child assessment and developmental screening	73.8
Materials and learning environment	78.9	Materials and learning environment	72.5
Understand/improve scores on Early Childhood Environment Rating Scale-Revised (ECERS)/Family Child Care Environment Rating Scale (FCCERS)/ Infant/Toddler Environment Rating Scale (ITERS)	78.2	Child behavior management	72.5
Child behavior management	76.8	Health and safety	72.5
Least Reported			
Content area	%	Content area	%
Special needs or inclusion	43.7	A specific curriculum	41.3
Relationship-based practices with infants and toddlers	29.6	Relationship-based practices with infants and toddlers	30.0
Licensing issues	25.4	Licensing issues	30.0
Accreditation	17.6	Accreditation	20.0
Business practices, program management, and/or fiscal management	13.4	Business practices, program management, and/or fiscal management	13.8
Number of respondents	142		81

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.11.

Other Characteristics of Coaching or Mentoring Support

Most staff reported receiving coaching or mentoring support at their center. The most frequently reported reasons for participation were center requirements and desire for self-improvement.

The vast majority of staff (90 percent in total) reported receiving coaching or mentoring support at their center. More than one third of staff reported receiving coaching or mentoring in person offsite (37 percent in total) and/or receiving coaching via an online, e-mail, or video mechanism (38 percent in total). Appendix exhibit 5B.12 provides response rates for all coaching locations.

The two most frequently reported reasons for participation in coaching or mentoring were a center participation requirement (50 percent in total) and a wish to improve their practice (self-improvement, 40 percent in total). Staff also reported financial stipends (28 percent) and free classroom materials (27 percent) as incentives for participating in coaching activities. About one quarter (24 percent) of staff reported that they received no incentives for participating in these

activities. Responses were similar among lead and assistant teachers. See appendix exhibit 5B.12 for additional details.

Finally, we asked staff if the coaching support they had received was provided by a specific program. Of our list of statewide programs, staff most frequently selected AB212/CARES Plus, Center on the Social and Emotional Foundations for Early Learning (CSEFEL), and Head Start coaches as the program through which they had received coaching support (exhibit 5.4).[12] About 19 percent of responding staff indicated they did not know what program provided their coaching support.

Exhibit 5.4. Statewide Program Providing Coaching Support by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
		Percentage	
AB212 or CARES Plus program	27.8	25.3	26.9
Center on the Social and Emotional Foundations for Early Learning (CSEFEL)	28.6	21.3	26.0
Head Start coaches	21.8	32.0	25.5
California Preschool Instructional Network (CPIN) coaches and on-site training and technical assistance	18.8	10.7	15.9
My Teaching Partner	14.3	13.3	13.9
Partners for Quality (PITC)	14.3	9.3	12.5
CA Early Childhood Mentor program	15.0	5.3	11.5
Child Signature program (CSP)	13.5	8.0	11.5
CA Inclusion and Behavior Consultation Network	4.5	2.7	3.9
Don't know/uncertain	15.0	25.3	18.8
Number of respondents	142	81	223

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.13.

NOTE: Respondents could select more than one program.

Family Child Care Respondents

Most FCCH respondents reported that they had received coaching or mentoring support, with self-improvement being the most frequently reported incentive for participation.

Among respondents who work in FCCHs, we observed somewhat different response patterns. However, our small FCCH sample size of 27 staff and lack of FCCH representation from all Consortia means there is little value of direct comparisons between center and FCCH groups.

[12] Appendix exhibit 5B.13 provides the full list of programs indicated by respondents, including those specific to an RTT-ELC Consortium. Response rates for Consortia-specific programs will be affected by the number of survey respondents within a given Consortium, which is why we focus on statewide programs in this chapter.

Among the 23 FCCH respondents in our sample who indicated received coaching or mentoring support, staff reported receiving an average of 53 hours from June 2014 through March 2015.

When asked about reasons for participating in coaching or mentoring activities, 18 of 23 FCCH staff (78 percent) reported that they wanted to participate for their own self-improvement. Nearly half (48 percent) of FCCH staff cited free classroom materials as an incentive for participation. Appendix exhibits 54B.14–5B.19 provide additional details on survey responses for FCCH staff. Direct contact with a coach in their home who understood the context in which they operated was highlighted in interviews with family child care providers. "When you're a daycare provider, you're in your house almost 24/7," explained one provider. "You're in your house, and there are very few people you can talk to in depth about human development, or child development, or these intricate little individuals running around your house. And with the coaches, their purpose is to come there and talk with you and see where you are and help you through certain challenges or revel with you in certain little victories you may have. … As a provider, you're giving so much energy continuously—to the children in your day care, to the parents, to your own family. And you come virtually last, if you even have time for that. So to actively be able to talk with someone who has the same passion or same interest that you do, it's refreshing. It's inspiring as well."

Noncredit Workshops or Training

Dosage

Most center staff reported participating in noncredit workshops or training, although reported monthly dosage was not necessarily consistent.

Most center staff reported that they participated in noncredit workshops or training at some point between June 2014 and March 2015, although not necessarily consistently over all months. Nearly three quarters (73 percent) of lead and assistant teachers reported that they had participated in at least some workshops or training during that period. This is fairly consistent with the percentage of staff (76 percent) who reported participating in workshops or training the *previous* year (June 2013 through May 2014). See appendix exhibits 5B.4 and 5B.6 for details. When considering the percentage of staff who participated in workshops or training at any point during the combined two-year period (June 2013 through March 2015), the total rises to 83 percent.

Of staff who participated in workshops or training from June 2014 to March 2015, the average total hours across that time period was 28 hours for both lead and assistant teachers (exhibit 5.5). The average *monthly* number of hours spent in workshops or training was 2.8. The hours that staff spent in workshops or training was reported to be lower in June and July. Among lead teachers, the highest average hours occurred in August (4 hours); assistant teachers reported the highest average hours in September (4.4 hours).

Exhibit 5.5. Noncredit Workshops or Training Hours by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
Total average hours from June 2014 to March 2015	27.8	28.0	27.9
Average hours per month	2.8	2.8	2.8
Number of respondents	132	69	201

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.20.

Content Areas

Although almost all staff reported spending some of their workshop or training time on all three core content areas, the most hours were spent on social and emotional development.

Almost all staff reported that some of their workshop or training time was spent on the three core content areas—language development/literacy, math/cognitive development, and social and emotional development—but that most workshop or training hours were spent on social and emotional development. Nearly all (98 percent) staff reported spending some workshop or training time on social and emotional development, 97 percent on language development/literacy, and 94 percent on math/cognitive development. Much of that time was devoted to social and emotional development. Half (49 percent) of staff reported that half or more of their workshop or training time was spent on social and emotional development while only 25 percent of staff reported that half or more of workshops or training time was spent on math. See appendix exhibit 5B.22 for details.

Of the 20 *other* potential content areas addressed through workshops or training, child assessment and development screening, teacher-child interactions, child behavior management, and materials and learning environment were reported as the top four content areas by both lead and assistant teachers (exhibit 5.6). For lead teachers, the fifth top content area covered by workshops or training was classroom management, whereas for assistant teachers it was a tie between health and safety and understanding/improving scores on the Early Childhood Environment Rating Scale (ECERS). The content areas *least* reported by all staff were business practices, accreditation, and licensing issues. Appendix exhibit 5B.23 provides the response rates for the full list of content areas.

Exhibit 5.6. Top 5 Most and Least Reported Noncredit Workshops or Training Content Areas by Staff Type: Centers (June 2014–March 2015)

Lead		Assistant	
Most Reported			
Content area	%	Content area	%
Child assessment and developmental screening	73.0	Child assessment and developmental screening	68.7
Teacher-child interactions	69.1	Teacher-child interactions	67.2
Child behavior management	64.3	Child behavior management	67.2
Materials and learning environment	61.9	Materials and learning environment	62.7
Classroom management	61.1	Health and safety; understand and improve scores on ECERS/FCCERS/ITERS	61.2
Least Reported			
Content area	%	Content area	%
A specific curriculum	34.9	Special needs or inclusion	38.8
Relationship-based practices with infants and toddlers	27.8	Relationship-based practices with infants and toddlers	32.8
Licensing issues	22.2	Licensing issues	14.9
Accreditation	11.9	Accreditation	13.4
Business practices, program management, and/or fiscal management	8.7	Business practices, program management, and/or fiscal management	9.0
Number of respondents	132		69

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.23.

Other Characteristics of Noncredit Workshops or Training

Most staff reported attending noncredit workshops or training in person. Top reasons for participation were self-improvement and center requirements.

The majority of staff reported participating in noncredit workshops or training in person—either offsite or at their centers. Overall, 70 percent of staff reported receiving in-person off-site training and two thirds of staff (67 percent) reported receiving training at their center. In addition, almost one third of lead teachers (32 percent) and one quarter (23 percent) of assistant teachers reported participating in remote trainings/workshops (i.e., online, e-mail, or video trainings). See appendix exhibit5B.24 for details.

When asked about whether there were any incentives or requirements to participate in workshops or training, the two most frequently reported by staff overall were that they wanted to participate for their own self-improvement (48 percent) and that the workshops or training were required by their center (also 48 percent). Finances influenced participation as well, with 40 percent of staff reporting that workshops or training provided for free were an incentive and 29 percent reporting that a financial stipend was an incentive for participating. Twenty percent of staff reported that they received no incentives for participating in workshops or training. See appendix exhibit 5B.24 for additional details.

When asked about which statewide program(s) had provided the workshops or training in which they participated, staff most frequently selected AB212/CARES Plus (36 percent), Desired Results Developmental Profile (DRDP) field training (31 percent), and CSEFEL (31 percent) (exhibit 5.7).[13] The CPIN also was reported by 29 percent of lead teachers. More than one in five (22 percent) responding staff indicated they did not know which program provided their workshops or training support, a similar rate of uncertainty as reported for coaching support programs.

Exhibit 5.7. Statewide Program Providing Noncredit Workshops or Training by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
		Percentage	
AB212 or CARES Plus program	35.8	36.7	36.1
Desired Results Developmental Profile field training	31.7	30.0	31.2
Center on the Social and Emotional Foundations for Early Learning /Teaching Pyramid	33.3	25.0	30.6
California Preschool Instructional Network (CPIN)	29.3	13.3	24.0
English language learners support	19.5	11.7	16.9
The Program for Infant/Toddler Care (PITC)	10.6	5.0	8.7
Family Child Care at Its Best	4.9	6.7	5.5
Don't know/uncertain	23.6	20.0	22.4
Number of respondents	132	69	201

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.25.

NOTE: AB212 or CARES Plus program includes county-specific programs such as ASPIRE, PIECES, and so on. Respondents could select more than one program.

Family Child Care Respondents

The majority of FCCH respondents reported that they had participated in workshops or trainings. Self-improvement, free participation, and free classroom materials were the top reported incentives for participation.

Among respondents from FCCHs, response patterns differed slightly from overall center responses. However, our limited FCCH sample size limits the value of direct comparisons between center and FCCH groups. That said, the 23 FCCH staff who indicated participating in workshops or trainings reported an average of 34 hours from June 2014 through March 2015.

Overall, many of the 23 FCCH staff reported participating in workshops or training that covered the three core content areas—social and emotional development (91 percent), language development/literacy (87 percent), and math/cognitive development (82 percent). More than two

[13] Appendix exhibit 5B.25 provides the full list of programs indicated by respondents, including those specific to an RTT-ELC Consortium.

thirds of FCCH staff reported that they participated in workshops or training for their own self-improvement (16 of 23 staff, 70 percent), more than half received a workshop or training for free (12 of 23 staff, 52 percent), and more than one third received free classroom materials (8 of 23 staff, 35 percent). Only 2 of 23 staff (9 percent) reported receiving a financial incentive. Appendix exhibits 5B.26-5B.31 provide additional details on survey responses for FCCH staff.

Peer Support

Dosage

A majority of center staff reported participation in formal peer support activities, although reported monthly dosage was not necessarily consistent.

Our questions about peer support attempted to characterize a type of QI activity that is more diverse and harder to define than most. On the survey, we asked respondents to tell us if they had participated in *formal* peer support arrangements as defined above. Beyond this definition, we know little about the variety of ways in which peer support is delivered. However, a small number of preliminary interviews that we conducted in the course of survey development provide some clues about this variation. For example, one interviewee indicated that in her center, peer support occurs on a regular basis in the context of staff meetings. During these meetings, staff are encouraged to address a topic together and share their experiences and insights. Another interviewee who is a family child care provider noted that she belongs to a family child care provider group in her area. One of the highlights of these meetings for her is the group's focus on issues concerning classroom practice during which members discuss ways in which they have attempted to address it.

More than half of center staff reported that they participated in formal peer support activities such as learning communities, peer support networks, or reciprocal peer coaching at some point from June 2014 through March 2015. Fifty-seven percent of all teachers reported that they had received at least some peer support during that period. This is fairly consistent with the percentage of staff (58 percent) who reported receiving peer support the *previous* year (June 2013 through May 2014). See appendix exhibits 5B.4 and 5B.6 for details. When considering the percentage of staff who received peer support at any point during the combined two-year period (June 2013 through March 2015), the total rises to 64 percent.

Among staff who participated in peer support activities between June 2014 and March 2015, the average total hours per staff person across that time period were 22.8 hours (exhibit 5.8), and the average *monthly* hours were 2.3. Lead teachers reported 21.4 average total hours per person and assistant teachers reported 25.2 hours. The reported hours spent in peer support activities were lower in the summer months (June, July, and August) than during the school year, similar to the pattern found for coaching support. See appendix exhibit 5B.32 for a monthly breakdown of the average number of peer support hours. When examining consistency of peer support received over time, we found that nearly half (45 percent) of total center staff reported at least some peer support during every month between September 2014 and March 2015 (see appendix exhibit 5B.34).

Exhibit 5.8. Formal Peer Support Hours and Frequency per Person by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
Hours			
Total average hours from June 2014 to March 2015	21.4	25.2	22.8
Average hours per month	2.1	2.5	2.3
Frequency			
Total average frequency from June 2014 to March 2015	12.8	11.5	12.4
Average frequency per month	1.3	1.2	1.2
Number of respondents	98	58	156

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibits 5B.32-5B.33.

Staff reported that, on average, they received peer support 1.2 times per month and a total of 12.4 times across all 10 months between June 2014 and March 2015 (exhibit 5.8).[14] During each month of the school year, 14 percent to 20 percent of teachers reported receiving peer support three or more times; these percentages are lower for the summer months. In fact, for June and July, the majority of center staff reported that they did not participate in peer support activities at all. See appendix exhibit 5B.33 for frequency as reported by month.

Content Areas

Of the three core content areas, the highest percentage of peer support hours was spent on social and emotional development and the lowest on mathematics. However, almost all staff received some peer support in all three core content areas.

Similar to the patterns for content areas reported for coaching and mentoring as well as workshops or training, nearly all staff reported receiving peer support in the three core content areas. More than 90 percent of staff reported spending time on social and emotional development, language development/literacy, and math/cognitive development (97 percent, 95 percent, and 93 percent, respectively). As was true with coaching and workshops, the highest percentage of peer support hours was spent on social and emotional development and the lowest on math. Thirty-nine percent reported that half or more of peer support time was spent on social and emotional development while only 25 percent reported that half or more of peer support time was spent on math. See appendix exhibit 5B.35 for details.

Of the 20 *other* potential content areas addressed through peer support activities, child assessment and development screening, teacher-child interactions, and child behavior

[14] Frequency of formal peer support was calculated by averaging assigned midpoint values to categories asked in the question. Appendix exhibit 5B.33 shows percentages of respondents for each category. Categories of frequency (and corresponding midpoint values) included: "Not at all" = 0, "1 or 2 times" = 1.5, "3 or 4 times" = 3.5, "5 or more times" = 5.

management were reported among the top five content areas by both lead and assistant teachers (exhibit 5.9). The other top five peer support content areas reported by lead teachers were materials and learning environment as well as understanding and improving CLASS. For assistant teachers, the other top five content areas included subjects other than language or math and physical development and health. The content areas *least* reported by all staff were business practices, accreditation, and licensing issues. Appendix exhibit 5B.36 provides the response rates for the full list of content areas.

Exhibit 5.9. Top Five Most and Least Reported Formal Peer Support Content Areas by Staff Type: Centers (June 2014–March 2015)

Lead		Assistant	
Most Reported			
Content area	%	Content area	%
Child assessment and developmental screening	69.2	Child assessment and developmental screening	69.6
Teacher-child interactions	67.0	Teacher-child interactions	66.1
Materials and learning environment	66.0	Subjects other than language or math	64.3
Understand/improve scores on CLASS	62.8	Child behavior management	62.5
Child behavior management	61.7	Physical development and health	60.7
Least Reported			
Content area	%	Content area	%
California Preschool Learning Foundations and Frameworks	40.4	Special needs or inclusion	42.9
Relationship-based practices with infants and toddlers	23.4	Relationship-based practices with infants and toddlers	39.3
Licensing issues	22.3	Licensing issues	23.2
Accreditation	10.6	Business practices, program management, and/or fiscal management	17.9
Business practices, program management, and/or fiscal management	8.5	Accreditation	12.5
Number of respondents	98		58

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.36.

Other Characteristics of Formal Peer Support

Most staff reported receiving peer support onsite at their center. Incentives for participation included self-improvement and center requirements, but several staff reported having no specified incentives for peer support participation.

Most staff reported receiving peer support in person on site at their center (87 percent); far fewer reported in-person support in other locations (38 percent). About one quarter (24 percent) of staff reported that they had received some peer support online, or via e-mail or video. Fewer than one out of 10 teachers (8 percent overall) reported that they had received peer support over the phone. See appendix exhibit 5B.37 for details.

The top two incentives or requirements to participate in peer support, reported by more than one third of respondents, included wanting to participate for their own self-improvement (37 percent) and being required to participate by their center (35 percent). Receipt of resources, either a stipend or free classroom materials, were less common—only 18 percent of staff reported these as reasons for participating in peer support activities. About one third (36 percent) of staff reported having no specified incentives for participating in peer support activities, a higher percentage than found for coaching or workshops participation. See appendix exhibit 5B.37 for additional details.

The program sponsorship of these formal peer support activities is less clear than for other QI activities. Only 25 percent of staff indicated that support was provided through a specific program; about the same percentage (26 percent) said that it was not provided by a program. Almost half (49 percent) said that they did not know whether the peer support they had received was offered through a particular program or not. See appendix exhibit 5B.38 for additional details.

Family Child Care Respondents

A minority of FCCH staff reported participating in any peer support activities. Of those who did participate, self-improvement was the most frequently reported incentive.

Very few FCCH staff in our survey reported participating in any peer support activities, so the following information is intended merely to highlight a few reported results. Among the 13 FCCH respondents, staff reported receiving an average number of 30 hours of peer support over the 10-month period between June 2014 and March 2015.

The most reported content areas for peer support included language development/literacy, social and emotional development, and math/cognitive development (all reported by 12 of 13 staff, 92 percent). Three other content areas also were reported by 83 percent of FCCH staff—materials and learning environment, child assessment and developmental screening, and family engagement. More than two thirds of FCCH staff (9 of 13, 69 percent) reported participating in formal peer support activities for their own self-improvement. Appendix exhibits 5B.39–5B.45 provide additional details on survey responses for FCCH staff.

Credit-bearing College or University Courses

Dosage

Some center staff reported participating in credit-bearing coursework at a college or university. For these staff, the average amount of coursework completed was nine semester-equivalent units.

Just over one quarter (27 percent) of center staff reported that they had participated in credit-bearing coursework at a college or university between summer 2014 and winter 2015. Credit-bearing coursework was reported by almost one quarter of lead teachers (24 percent) and almost one third of assistant teachers (32 percent). This is a slight decrease from the percentage of staff (28 percent of lead teachers and 37 percent of assistant teachers) who reported participating in

credit-bearing courses the *previous* year (summer 2013 through winter/spring 2014). See appendix exhibits 5B.4 and 5B.6 for details. When considering the percentage of staff who reported completing credit-bearing courses at any point during the combined two-year period (summer 2013 through winter 2015), the total increases to 41 percent of all staff. These rates of reported college course-taking may relate to the fact that many teaching staff, especially assistant teachers, do not yet have a college degree and may be working toward one as part of professional development.

For the staff who reported completing credit-bearing courses, the average number of semester-equivalent units completed per person from summer 2014 through winter 2015 was nine units (exhibit 5.10). Lead teachers reported about eight units per person, on average, and assistant teachers reported about 10 units. For both lead and assistant teachers, the reported average units completed was highest during the fall semester (compared with the summer or winter semesters).

Exhibit 5.10. Average Credit-bearing Course Units per Person by Staff Type: Centers (June 2014–March 2015)

	Lead	Assistant	All
Total average units from summer 2014 to winter 2015	7.7	10.3	8.9
Average units per semester	2.6	3.4	3.0
Number of respondents	42	32	74

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.46.

NOTE: Units for quarter-system courses are converted to semester-equivalent units.

Content Areas

Most of the staff taking credit-bearing courses took courses with a focus in early childhood education.

Overall, the majority of staff who were taking credit-bearing courses reported that some or all of these courses had an early childhood education (ECE) focus. Almost three fourths (74 percent) of lead teachers and more than half (56 percent) of assistant teachers reported that their credit-bearing courses were exclusively ECE-focused. See appendix exhibit 5B.47 for details.

For the teachers who indicated taking ECE coursework and noted their topic areas, about half (51 percent) reported that this coursework included a focus on child growth and development (see appendix exhibit 5B.47). Just under one third of teachers (30 percent) reported that their coursework included the topics of child/family and community, introduction to curriculum, and/or observation and assessment. Overall, fewer than one in 10 teachers reported that they had taken coursework covering adult supervision (9 percent), a practicum (8 percent), or administration (2 percent).

Other Characteristics of Credit-bearing Courses

Most staff reported attending courses on campus, and some staff reported receiving supports such as financial stipends or academic counseling for their participation in credit-bearing coursework.

Most staff reported that they had attended courses in person on a college or university campus. Overall, 85 percent of teachers reported taking courses on campus while 21 percent reported that they had taken a college or university course online. Only 9 percent reported taking a course in person off campus. See appendix exhibit 5B.48 for details.

We also asked teachers about the types of support they had received for their participation in credit-bearing courses. Almost one third (32 percent) of teachers indicated that they had received a financial stipend or scholarship for their participation in credit-bearing courses. Among types of nonfinancial support, one fourth (25 percent) of staff stated that they had received academic counseling or advisement, and almost one fifth (19 percent) reported receiving tutoring to support their participation in credit-bearing courses. Overall, 48 percent of staff who participated in credit-bearing courses reported that they had not received any nonfinancial support to take courses. For additional details, see appendix exhibit 5B.49.

A smaller subset of center staff (41 total) reported on the statewide program(s) providing financial or nonfinancial support for credit-bearing courses, and staff most frequently selected AB212 or CARES Plus (22 of 41 staff, 54 percent) (exhibit 5.11).[15] One quarter or more of the 20 lead teachers also reported receiving support for credit-bearing courses through the Child Development Training Consortium (CDTC) (30 percent) or a Child Development Grant through the Student Aid Commission (25 percent). More than one third of assistant teachers (8 of 21, 38 percent) and one in 10 lead teachers (2 of 20, 10 percent) reported that they did not know which program had provided support for their credit-bearing course participation.

Exhibit 5.11. Statewide Programs Providing Support for Credit-bearing Courses by Staff Type: Centers (June 2014–March 2015)

Program	Lead	Assistant	All
AB212 or CARES Plus program	60.0	47.6	53.7
Child Development Training Consortium	30.0	14.3	22.0
Child Development Grant (Student Aid Commission)	25.0	4.8	14.6
Family Child Care at Its Best	5.0	0.0	2.4
Child Signature Program (CSP)	5.0	0.0	2.4
Don't know/uncertain	10.0	38.1	24.4
Number of respondents	20	21	41

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.50.

NOTE: AB212 or CARES Plus program includes county-specific programs such as ASPIRE, PIECES, and so on. Respondents could select more than one program.

[15] Appendix exhibit 5B.50 provides the full list of programs indicated by respondents, including those specific to an RTT-ELC Consortium.

Family Child Care Respondents

A small number of FCCH staff reported participating in credit-bearing courses. Of those that did, more than half had received a financial stipend or scholarship for their participation.

As in the section above on peer support activities, a very small number of FCCH staff in our survey reported participating in any credit-bearing courses, so the following information is intended merely to highlight a few reported results. Among the 13 FCCH respondents who participated in credit-bearing coursework, 62 percent (8 of 13 staff) reported attending courses on campus. Almost half (46 percent) of the 13 FCCH staff reported taking courses online. See appendix exhibit 5B.53 for this information.

More than half of FCCH staff (54 percent, or 7 of 13 staff) reported receiving a financial stipend or scholarship for their credit-bearing courses, and one third (4 of 12) of FCCH staff reported receiving academic counseling or advisement. Nine FCCH staff reported on the statewide program(s) providing supports for their participation in courses, and although one third (or 3 of 9 staff) reported that they did not know what program provided these supports, two of nine FCCH staff indicated that supports came from the CDTC and one of nine staff identified AB212 or CARES Plus as the program. See appendix exhibits 5B.51–5B.55 for additional details.

Financial Incentives

A minority of staff reported receiving a financial incentive to promote their participation in QI efforts. Among those receiving a financial incentive, the average dollar amount was more than $1,000 per person.

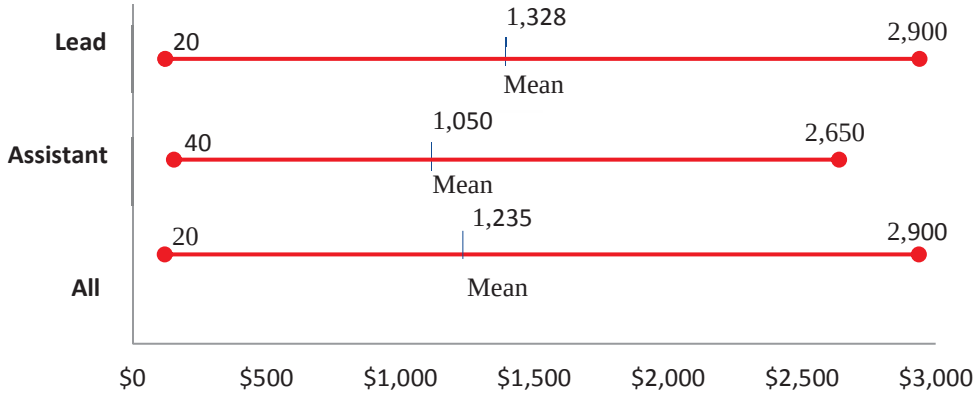
Fewer than half of staff reported having received any financial incentive, such as a scholarship or stipend, to promote their participation in QI efforts between July 2013 and June 2015. About one third of staff (33 percent) reported receiving some type of financial incentive in the most recent fiscal year between July 2014 and June 2015. [16] This is fairly consistent with the percentage of staff (32 percent) reporting that they had received a financial incentive in the previous year (July 2013–June 2014). See appendix exhibits 5B.6 and 5B.56 for these data. When the two years are combined, a total of 42 percent of staff said they received some financial incentive between July 2013 and June 2015.

The value of financial incentives received was more than $1,000 per person on average. Among those receiving a financial incentive in any amount, the average amount was $1,235 (exhibit 5.12). The mean dollar amount of financial incentives for lead teachers was $1,328, with a range of $20 to $2,900. Among assistant teachers, the mean dollar amount was $1,050 and ranged from $40 to $2,650. To place these amounts in context, the 2014 average annual wage for California

[16] Because staff were surveyed before the end of June 2015, they were asked to indicate what they had received by the time of the survey or what they expected to receive through June 2015.

preschool teachers reported by the Bureau of Labor Statistics was $34,070, and the average wage for other nonpreschool child care workers was $24,810.[17]

Exhibit 5.12. Mean and Range of Financial Incentive Dollar Amount Received by Staff Type: Centers (July 2014–June 2015)



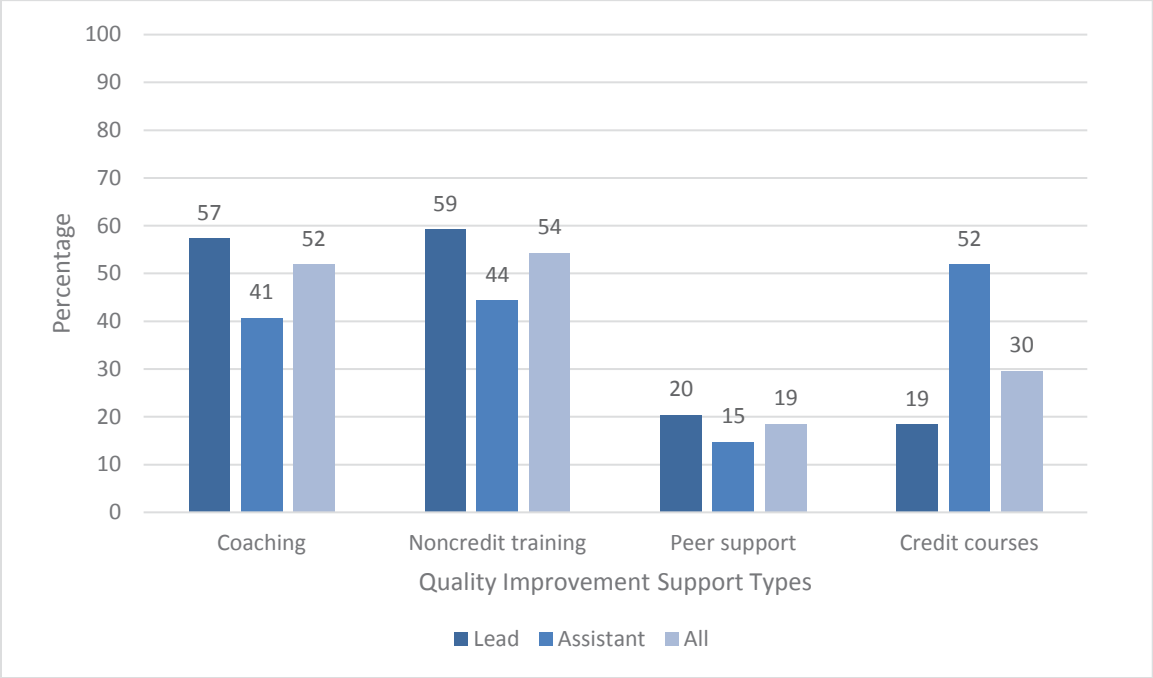
SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.57.

NOTE: Dollar amounts are calculated based on nonmissing cases for the 59 lead and 31 assistant teachers who reported having received any financial incentives between July 2014 and June 2015.

When asked to indicate which QI efforts their financial incentives covered, the majority noted coaching and mentoring and noncredit workshops or training (exhibit 5.13). The majority of lead teachers reported that financial incentives covered coaching or mentoring (57 percent) and noncredit workshops or training (59 percent), and fewer lead teachers reported financial incentives for credit-bearing courses (19 percent) or peer support activities (20 percent). For assistant teachers, credit-bearing courses were the most common type of QI covered by financial incentives (52 percent), with noncredit workshops or training and coaching following at 44 percent and 41 percent. This pattern was slightly different for the prior year (July 2013–June 2014), in which 57 percent of assistant teachers reported that their financial incentives covered noncredit workshops or training and 37 percent reported that they covered credit-bearing courses.

[17] See "May 2014 State Occupational Employment and Wage Estimates: California" at http://www.bls.gov/oes/current/oes_ca.htm.

Exhibit 5.13. Types of Quality Improvement Covered by Financial Incentives: Centers (July 2014–June 2015)



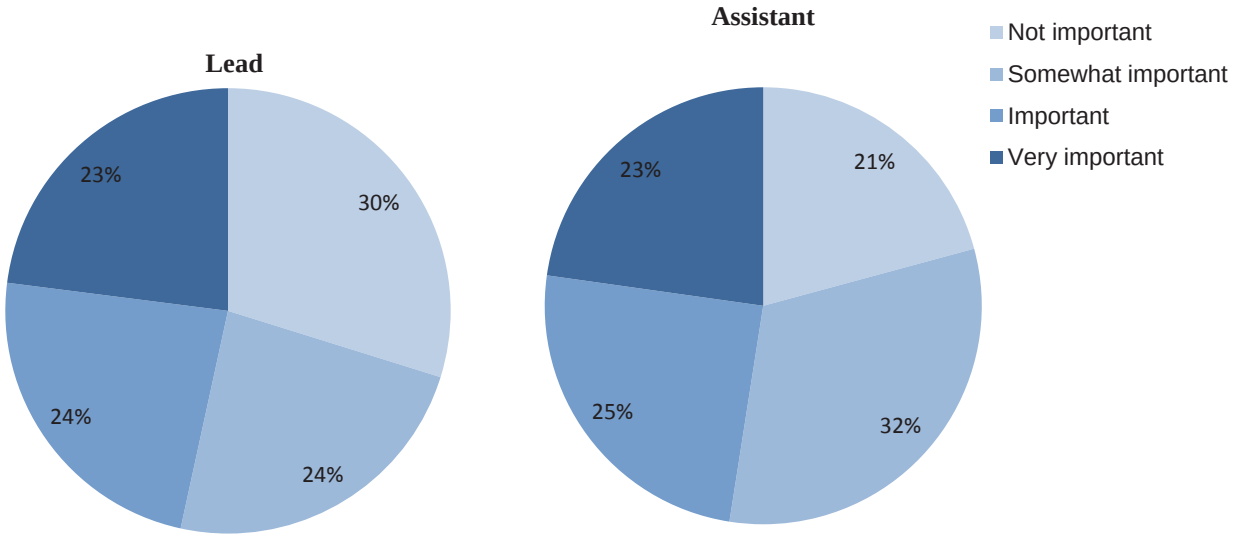
SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.58.

NOTE: Percentages are calculated based on nonmissing cases for the 59 lead and 31 assistant teachers who reported having received any financial incentives between July 2014 and June 2015.

Almost half (48 percent) of staff who reported receiving a financial incentive identified AB212 or CARES Plus as the program that provided the incentives between July 2014 and June 2015 (up from 31 percent in the previous year). All other statewide programs (e.g., CSP, Child Development Grants) were selected by less than 10 percent of staff as financial incentive providers. One exception to this was the Career Incentive Grants provided by the CDTC, which was selected by 17 percent of assistant teachers. Very few staff (2 percent) reported being uncertain about the provider of their financial incentive. Appendix exhibit 4B.59 includes response rates for all financial incentive programs.

The majority of staff (74 percent in total) indicated that the availability of financial incentives was at least somewhat important in their decision to participate in QI efforts (exhibit 5.14), with 23 percent of lead and assistant teachers alike indicating that financial incentives were very important in their decision. At the same time, 30 percent of lead teachers and 21 percent of assistant teachers indicated that financial incentives were not at all important in their decision to participate.

Exhibit 5.14. Percentage of Staff Indicating Financial Incentives Were Important to Their Decision to Participate in QI: Centers (July 2014–June 2015)



SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.56.

NOTE: Percentages are calculated based on nonmissing cases in the sample of 174 lead and 105 assistant teachers. Totals may not sum to 100 due to rounding.

Family Child Care Respondents

Though the majority of FCCH staff indicated that financial incentives were at least somewhat important in their decision to participate in QI efforts, fewer than half reported that they actually had received such incentives.

With the caveats discussed earlier, the following information provides several descriptive highlights of results but should be interpreted with caution. Among the 27 FCCH respondents in our sample, 41 percent (or 11 of 27) reported receiving some type of financial incentive in the most recent fiscal year between July 2014 and June 2015. Overall, 21 of 25 FCCH staff (84 percent) reported that financial incentives were at least somewhat important in their decision to participate in QI efforts.

Among the 11 FCCH respondents who reported receiving financial incentives in any amount between July 2014 and June 2015, the average amount reported was $972, with a range of $50–$1,500. When asked about which QI types were covered by these financial incentives, credit-bearing courses were the most reported (46 percent, or 5 of 11 staff), followed by coaching (36 percent, or 4 of 11 staff). Six of 11 FCCH staff (55 percent) identified AB212 or CARES Plus as the program providing their financial incentives. See appendix exhibits 5B.60–5B.63 for additional details on FCCH staff responses.

Perceptions of Quality Improvement Activities

Identifying QI Activities

Most center staff learned about QI activities through their program or program director.

The majority of center staff (76 percent) reported that they learn about QI activities through their program or program director, which is by far the most reported source. About one fourth (26 percent) reported learning about QI opportunities through First 5 California, and just over one fifth said that they learned about QI activities through their colleagues (22 percent) or their County Office of Education (21 percent). Fewer staff reported learning about QI activities through their own research (13 percent), their local QRIS (10 percent), or their local county-level First 5 (5 percent). No center staff reported that they learned of QI activities through their local Resource and Referral (R&R) agency. Patterns are similar among lead and assistant teachers. See appendix exhibit 5B.64 for additional details.

Reasons for and Barriers to Quality Improvement Participation

Personal interest, QI plans, and supervisor recommendations were among the top reasons for staff participating in QI activities. Many staff reported that not having enough time was a barrier to their participation in QI activities.

The top reasons for participating in QI activities were personal interest, QI plans, and supervisor recommendations. Overall, 60 percent of center staff reported that they participated in QI activities because of their own personal interest in a given topic or activity. “I can see the change in my own teaching and that made a change in the students,” noted one classroom teacher in an interview. “I believe children deserve a quality program … it does make a difference when the teachers are active in workshops or in school. We’re learning new things all the time.” Another teacher stated, “We do it for the personal reward. The staff continues their education because we can always get better.” Half (48 percent) of the survey respondents stated that their decision to participate in QI activities was because they were identified as part of their classroom or site QI plan. Supervisor recommendation was cited by 42 percent of staff as a reason for participation. Overall, only 12 percent of staff indicated that the offer of financial incentives was a factor in their decision to participate. Although this result seems to contradict the data reported above in Exhibit 5.14 that just under half of staff consider incentives important or very important in their decision to participate in QI efforts, the inconsistency may be more apparent than real. To a group of people whose earnings are low, receiving a financial incentive is likely to be important always. But valuing a financial incentive does not negate the fact that these are dedicated professionals who are motivated to participate to improve their practice. See appendix exhibit 5B.64 for additional details.

When asked about barriers to participating in QI activities, one third of all staff indicated there were no barriers preventing their participation. For those who did face barriers, time was the top challenge cited by both lead and assistant teachers. Slightly over half (54 percent) of staff said that they did not have enough time to participate in QI activities. One classroom teacher reported in an interview, “It is hard to find the time to get it all together. When we are working with children all day.” Another teacher spoke about the need for more online resources to improve

access. “The training and the workshops are only offered certain times and certain days … and I’m already scheduled that time during the year. I’m thinking resources that we’d have access to … a website where you could get more information and have access to it whenever you need it.”

Other barriers noted were cost (19 percent of staff said that that QI activities were too expensive) and distance (17 percent of staff said that activities were too far away or difficult to get to). One interviewed teacher indicated, “I wouldn’t mind paying to go back to school because in the end it will benefit me. But for the fact that we don’t get paid much—I can’t afford it.”

When asked questions about whether language may be a barrier to participating in QI activities, only 2 percent of staff expressed that activities not being provided in their primary language was a barrier and less than 1 percent of staff reported that they were not very or not at all comfortable with QI activities provided in English. See appendix exhibit 5B.64 for additional details.

Perceived Helpfulness of QI Activities

Most staff reported that each QI type had been helpful or very helpful, though coaching or mentoring was reported as most helpful in comparison to the other QI types.

When staff were asked how helpful they found each QI type (coaching or mentoring, workshops or training, credit-bearing courses, or formal peer support) to be for improving their practice with children, the majority reported that *each* QI type had been helpful or very helpful. At 89 percent, workshops or training had the highest helpfulness ratings. Of staff who had participated in coaching or mentoring, 86 percent found that it had been helpful or very helpful. One teacher explained during an interview, “It is good to get an outsider to help with the group …. It is nice to have someone to share with about what we are working on and what we are doing. It is reassuring.” Most staff also found credit-bearing courses (86 percent) or formal peer support (82 percent) to be helpful or very helpful. Among the four QI types, credit-bearing courses had the highest percentage (55 percent) of participants rating this type of QI support as *very* helpful. See appendix exhibit 5B.65 for additional information.

To assess how the different QI types compared with each other in rated helpfulness, we calculated comparative QI helpfulness ratings for those teachers who had participated in more than one type of QI and indicated which they found most helpful (see appendix exhibit 5B.66).[18] More than half of respondents (57 percent) reported coaching or mentoring as most helpful in comparison with one or more other QI types. Almost one third (31 percent) of staff reported that workshops or training were most helpful, and a similar percentage (31 percent) chose credit-bearing courses. Only 8 percent of staff reported formal peer support as the most helpful QI activity in comparison with the other types.

---

[18] If teachers only participated in one QI type, they were not asked to compare that QI type to any other QI types. If teachers participated in two, three, or four of the QI types from June 2014 through March 2015, they were asked to choose the most helpful among the two, three, or four QI types they had participated in. Among all center staff, the following percentage of staff participated in a given QI type, in addition to at least one of the other three types: coaching or mentoring (75 percent), noncredit workshops (69 percent), formal peer support (56 percent), and credit-bearing courses (25 percent).

Topic Areas Needing More Support

Child behavior management was the topic area on which most center staff would like more support or training.

The most mentioned topic area on which center staff would like more support or training to improve their practices or achieve career goals is child behavior management (exhibit 5.15 presents the top 11 topics reported). Overall, 60 percent of staff reported that they would like additional behavior management support, with very similar rates for lead and assistant teachers. Almost half of staff (48 percent) indicated they would like additional training on language development/literacy, and 42 percent would like additional training on social and emotional development as well as math/cognitive development. An equal percentage (42 percent) would also like more support or training on special needs or inclusion. See appendix exhibit 5B.67 for additional information on all topic areas reported.

Exhibit 5.15. Most Reported Topic Areas Staff Would Like to Receive More Support or Training on by Staff Type: Centers

	Lead	Assistant	All
		Percentage	
Child behavior management	59.9	59.4	59.7
Language development/literacy	49.7	44.6	47.8
Social and emotional development	41.3	43.6	42.2
Special needs or inclusion	42.5	41.6	42.2
Math/cognitive development	41.9	41.6	41.8
Child assessment and developmental screening	35.9	45.5	39.6
Subjects other than language or math	34.1	36.6	35.1
Classroom management	33.5	36.6	34.7
Materials and learning environment	29.3	35.6	31.7
English language development	32.3	30.7	31.7
Family engagement	34.7	26.7	31.7
Number of respondents	174	105	279

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. For details, see appendix exhibit 5B.67.

NOTE: The top 11 topic areas are ordered from highest to lowest by total percentage reported for all staff.

Family Child Care Homes

All FCCH staff reported that the QI activities they participated in were helpful. However, some staff indicated that lack of time and language issues acted as barriers to participation.

Respondents who work in FCCHs expressed somewhat different perceptions of and experiences with QI. However, we again note the limited sample size (27 FCCH staff in total) with which to make comparisons or draw inferences. We note a few FCCH staff results below for descriptive purposes. See appendix exhibits 5B.68–5B.70 for further details.

Forty-one percent, or 11 of 27, FCCH staff said that they learned about QI activities through their colleagues, and about one quarter, or 7 of 27, reported that they learned of QI opportunities through their local R&R. The majority of FCCH respondents (22 of 27) reported that they decided to participate in QI activities because of their own personal interest in a given topic or activity. One family child care provider explained in an interview, "The more I started going back to college and participating in the offsite organizations and classes and groups, the more I realized that it was just a huge boost for me personally but also for the children I was educating in my program." The second highest-rated reason for participating in QI activities for FCCH staff was that a given activity was identified as part of their QI plan (10 of 27 FCCH staff). All FCCH staff reported that the QI types they had participated in were either helpful or very helpful.

As with center staff, FCCH staff cited lack of time as the biggest barrier to participation (50 percent, or 13 of 26 staff). One family child care provider interviewed who was currently working on completing a degree described, "Over recent time, a lot of the classes are trying to be done in the evenings, because family child care, we work the day. We don't have subs and can't get that kind of sub as often as the classes are, so the strategy of having classes at night and on weekends have really helped. But family child care, you work very long hours, so it's very hard." Some FCCH staff also reported that language issues created barriers to QI activity participation. Five of 26 staff reported that the activities they wanted to participate in were not provided in their primary language, and five FCCH staff also reported that they were not comfortable with QI activities provided in English.

Finally, FCCH staff identified many of the same topic areas as center staff as ones for which they would like to receive more support or training. These included language development/literacy (54 percent), social and emotional development (46 percent), and child behavior management (46 percent). See appendix exhibit 5B.70 for a full list of reported topic areas.

Summary

Nearly all of the teachers who responded to the staff survey reported engaging in QI activities between September 2014 and March 2015. It is notable that four out of five respondents committed time and resources to engage in activities with the aim to improve their knowledge and practice. Although there is some variation in the time commitment and mix of activities pursued across teacher type and care setting, assistant and lead teachers working in centers and staff working in FCCHs all reported substantial engagement in QI activities. Exhibit 5.16 shows the average hours reported for center staff for three types of QI activities (credit-bearing courses are not presented because they are reported as semester units and are not equally comparable to hours). Moreover, these activities appear to be a regular part of teachers' lives: many teachers reported consistent QI activity participation across months during the school year. We note again that these staff members work in sites that are voluntarily participating in the QRIS; it is not known if these same patterns of QI activities would hold for staff in non-QRIS sites. It also may be the case that QRIS staff members who chose to complete our survey participate in QI activities at different levels or in different ways than those who chose not to complete the survey. Thus, these findings represent QRIS survey respondents and are not necessarily generalizable to a wider population.

Exhibit 5.16. Average Hours per Person by QI Type: Centers (June 2014–March 2015)

	Coaching	Noncredit Training	Peer Support
Total average hours per person from June 2014–March 2015	22.2	27.9	22.8
Average hours per person per month	2.2	2.8	2.3
Number of respondents	223	201	156

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. Percentages reported in exhibits 5.1, 5.5, and 5.8.

Of the four activity types that they were asked about on the survey, teachers reported the most engagement with coaching and mentoring. They also found coaching and mentoring activities to be the most helpful when compared with activities in the three other categories: noncredit workshops or training, credit-bearing courses, and peer support activities. However, the teachers found all four QI types to be valuable: more than 80 percent rated each type as helpful.

Coaching and the other types of QI activities generally focus on three core content areas—language development/literacy, math/cognitive development, and social and emotional development. An emphasis was placed on social and emotional development as measured by hours spent in coaching and mentoring, workshops or training activities, and peer support activities. However, many other topics were discussed as well, such as teacher-child interactions and child assessment and developmental screenings; the coverage on these topics varied to some degree by teacher position and program type. Exhibit 5.17 provides a comparison by QI activity of the top 10 content areas reported by center staff. The content areas *least* reported by all staff across QI types were business practices, accreditation, and licensing issues.

Exhibit 5.17. Most Reported Content Areas by QI Type: Centers (June 2014–March 2015)

Content Area	Coaching	Noncredit Training	Peer Support
		Percentage	
Social and emotional development	98.2	97.5	96.7
Language development/literacy	96.4	97.4	94.6
Math/cognitive development	94.2	94.0	93.0
Teacher-child interactions	81.2	68.4	66.7
Understanding or improve scores on CLASS	77.1	60.1	58.7
Materials and learning environment	76.2	62.2	61.3
Child behavior management	74.9	65.3	62.0
Child assessment and developmental screening	74.0	71.5	69.3
Understand or improve scores on ECERS/FCCERS/ITERS	72.7	57.5	55.3
Classroom management	69.5	56.5	55.3
Number of respondents	223	201	156

SOURCE: Authors' analysis of the 2014–15 California QRIS Study Staff Survey. See appendix exhibits 5B.11, 5B.23, and 5B.36.

When asked to indicate topics in which they would like more support or training, the most mentioned topic area was child behavior management. The next most mentioned was language development/literacy, then additional training on social and emotional development, math/cognitive development, and more support or training on special needs or inclusion.

Finally, teachers reported being motivated to engage in QI activities for a variety of reasons, including their own self-improvement, as part of a classroom or site QI plan, or because it was suggested by a supervisor. Incentives such as free participation, a stipend, or free classroom materials also motivated participation to some degree. There were some differences in motivational patterns across QI types, but these patterns suggest that receiving external incentives is one of several motivating reasons for QI participation. One third of center staff reported no barriers to participating in QI activities.

These findings are encouraging in a number of ways. First, the data suggest that teachers consider QI to be important and valuable, and devote significant time to it on an ongoing basis. Moreover, they find coaching most valuable, which is consistent with research that generally finds that coaching can be effective in improving program quality in some circumstances. At the same time, research evidence has not yet determined the preferred dosage or other components of successful coaching models (AIR and RAND 2013). We do not know, for example, if frequency and total hours should be considered equally important, or if two hours per month is sufficient to achieve significant changes in teacher practices that affect child outcomes. Moreover, much of the extant research on informal trainings, such as noncredit workshops, suggests that effects are minimal or nonexistent when provided alone, though they may hold promise when combined with other QI supports such as coaching (AIR and RAND 2013). Thus, the survey findings presented here are useful as descriptive indicators of the QI activities in which this sample of QRIS staff participate, but more research evidence is needed to assess whether they make a difference in outcomes of interest.

Importantly, the nature of the QI in which teachers participate appears to support attainment of the standards required in California's QRIS [as described in the Hybrid Rating Matrix]. Based on our description of QI content, it appears that a substantial percentage of teachers are engaging in QI activities that address topics relevant to several elements measured in the RTT-ELC Hybrid Rating Matrix, such as improving CLASS or ERS scores. When teachers are asked what kinds of help they would like more of, their responses also are topics likely to address rating elements. Moreover, there is evidence in these data that state programs such as AB212 and CARES Plus are helping to support these efforts. Moving forward, it will be important to examine more specifically the variations in the content and quality of these QI offerings across consortia and their relationships with program quality and child outcomes.

Chapter 6. Program-Level Quality Improvement Supports

Key Findings

This chapter provides descriptive information about program-level supports for quality improvement (QI) activities from a survey of 93 directors or administrators representing a total of 102 sites participating in the study.

- Almost all centers had directors who participated in personal QI activities, including learning about social and emotional development, language development and literacy, and child assessment and developmental screening.
- About two fifths reported that their sites had received some financial benefits since enrolling in the RTT-ELC QRIS (though several directors reported that they did not know if the sites they supervised had received benefits).
- The majority of center directors indicated that a minimum number of professional development hours is required each year for staff. Two out of three centers required coaching for all staff, and a similar number required training for all staff. Most center directors indicated that coaching or mentoring is the most effective QI type.
- The majority of center directors reported observing classrooms to ensure new knowledge is implemented and checking in with staff to ensure they have the needed resources to do so.

The descriptive analyses are not representative of all programs in California, or of all programs participating in the California QRIS. They represent a self-selected sample of directors of programs participating in the California QRIS.

This chapter provides descriptive information about program-level support for quality improvement (QI) activities, including QI supports for those who oversee these programs, to address the following research questions:

- RQ 9. What are early learning staff's experiences with quality improvement activities?
- RQ 2. What incentives or compensation strategies are most effective in encouraging QRIS participation?

To collect program-level data, we conducted an online director survey for the 142 sites participating in the outcomes study to supplement the staff survey that we reported on in the previous chapter. By surveying program administrators, we can learn more about what supports programs offer their staff (and whether or not staff are aware of or take advantage of those activities). Examining supports from the program level also enables us to learn about how administrators perceive the QI activities that staff rated in the staff survey.

In designing this survey, we had to take into account the fact that some programs operate in more than one site, and a single administrator may oversee those multiple sites. Given this complexity, we asked those directors who administer more than one site to report separately for each site. This enabled us to examine site-level data and ensure that variations across sites are included in the analyses.

The results provided here describe survey responses from a select set of directors or administrators who oversee staff in centers or family child care homes (FCCHs) participating in the Race to the Top-Early Learning Challenge Quality Rating and Improvement System (RTT-ELC QRIS) within the study's focal Consortia.[19] We asked them about site-level QI activities within the center or FCCH they represent.[20] Findings reported in this chapter do not represent all of the sites included in the previous chapter, nor do they represent programs that have chosen not to participate in the RTT-ELC QRIS. As was the case in chapter 5, this chapter describes survey results and document response patterns among represented QRIS programs. It is important to note that, based on the descriptive nature of the survey, we are able to report differences in site-level QI activities but are not able to explain the reasons for the differences or analyze variations in the quality of the activities.

Analysis Approaches

- Descriptive analysis of early learning director survey responses
- Responses representative of sites rather than individual directors
- Report question response percentages and mean values as appropriate

Data and Sample

- In total, 142 directors representing center or FCCH sites participating in the outcomes study were surveyed.
- Respondents include 93 directors or administrators representing a total of 102 sites—89 centers and 13 family child care homes—across the 11 focal Consortia for a response rate of 72 percent.
- Sample includes sites with directors who are largely Hispanic or White (69 percent), speak English as their primary language (79 percent), and have a Bachelor's degree or above (73 percent).
- Sites have directors with an average of 15 years' experience working with young children.
- More than half of sites reported serving child populations in which a majority speaks a language other than English and/or receives free or subsidized care.

In addition to the survey of directors, we also conducted in-depth interviews with a small sample of early learning staff, including 13 center directors, five center teachers, and seven family child care providers. We draw on findings from these interviews throughout this chapter to add context to the survey responses presented here.

---

[19] All but two center respondents self-identified as the director, site supervisor, or other administrator. The others indicated an instructional support-focused job role.

[20] Not all invited directors responded to the survey invitation (see appendix 6B. for response rates). In a few cases, a director responded to this survey even though no teaching staff from the site responded to the staff survey. Moreover, programs that elected to participate in RTT-ELC may be different from programs that chose not to participate; they may differ in program characteristics and QI supports that may relate to their responses reported here.

Director and Site Sample Characteristics

In this chapter, we refer to center directors or administrators and family child care operators who responded to this survey as directors.[21] Of the 142 sites in our final survey sample, we received complete survey responses from directors representing a total of 102 sites or a 72 percent response rate—89 centers (87 percent of site responses) and 13 FCCHs (13 percent of site responses). The responses reported here represent the program site level as noted above. In some cases, directors oversee more than one site. In those cases, respondents provided information for each unique site. Indeed, 93 directors responded for 102 sites. To reflect this, we report on the sites in this chapter rather than on individual directors.[22] Responding directors represent all 11 focal Consortia (see appendix exhibit 6B.1).

Overall, the characteristics of our sample include sites with directors who are largely middle-aged, are Hispanic or White, and have a Bachelor's or Master's degree. The majority of directors reported that they were 40 to 59 years old (73 percent), whereas only 17 percent reported that they were 30–39 years old; 9 percent reported that they were 60 years or older. Just under half of the directors identified as Hispanic (46 percent), while almost one quarter identified themselves as White (23 percent). English was the primary language for the majority of directors (79 percent), with 18 percent reporting that their primary language was Spanish. Almost three quarters of directors (73 percent) reported that they had attained at least a Bachelor's degree; half of those also reported they had attained a Master's or doctoral degree. That said, almost one quarter of directors reported that they were *currently* enrolled in a college degree program (23 percent), with 20 percent reporting that their major was related to early childhood. Half of the directors in our survey sample have a Child Development Site Supervisor Permit, but only 40 percent have a Child Development Director Permit. Directors reported an average of 15 years of director or teaching experience with children age 0–5; the range of experience spanned from less than one year up to 38 years.

Based on director reports, the majority of sites represented in our director survey serve high percentages of students who speak a language other than English or who are receiving free or subsidized care. More than half of the site directors reported that 51 percent or more of their children speak a language other than English. Almost three quarters of the sites (71 percent) reported serving child populations in which more than 75 percent of children receive free or subsidized care.

According to directors, the majority of the sites (59 percent) represented in our director survey had received an RTT-ELC QRIS rating of Tier 3 or higher. Almost one third of sites were rated as Tier 4 (32 percent), while 15 percent of sites were Tier 3 and 12 percent were Tier 5.

---

[21] FCCH staff who were identified in our sample as "lead teachers" received questions from both the staff survey and the director survey in an effort to recognize their dual roles as both teachers and FCCH operators.

[22] For example, if a director supervises two sites in our survey sample, she was asked to complete survey responses for each site independently. For questions that pertain to her personal characteristics, such as education level, we apply her responses to each site. In this example, her education level would be attributed to two sites when we report on the director characteristics for all sites because she is the director for each of her two sites.

However, responding directors representing more than one quarter of sites in our sample did not know their current rating (27 percent), and another 9 percent reported they had not yet received an RTT-ELC rating. Appendix exhibits 6B.2 through 6B.4 include a detailed breakdown of these sample characteristics data.

FCCH respondents represent only 13 percent, or 13 sites in our sample. Their distribution across Consortia was uneven: 62 percent of all FCCH respondents work within two of the 11 Consortia, and four Consortia do not have any FCCH directors represented in the director survey. Given the small FCCH sample and the lack of representation across all Consortia, we combined the FCCH responses with center director responses in cases where all directors were asked the same question. However, some questions were asked only of center directors. Where a question was only asked of FCCH respondents, we report their responses but note, as we did in the previous chapter, that these FCCH survey results should be interpreted with caution. Note that appendix 6A provides the survey questionnaire, and appendix 6B includes complete exhibits of the data referenced in this chapter.[23]

RTT-ELC QRIS Participation

Reasons for QRIS Participation

Improving program quality was the most popular reason that led sites to participate in the RTT-ELC QRIS.

Respondents were asked to indicate which of nine reasons led their site to agree to participate in the RTT-ELC QRIS. The most popular reason was to improve program quality—as might be expected, more than three quarters of respondents (82 percent) indicated that their sites chose to participate for this reason. More than half (61 percent) of directors reported that their sites participate to gain new ideas, and almost half (48 percent) reported participating in the RTT-ELC QRIS to get the technical assistance that it offers. “The coaching for my staff – that was it,” one director noted when interviewed. “As a director, you have [new staff] who have the education but not the experience. Also you have the staff that have the experience but not the education. [I wanted to] help bring those two together.”

More than a third noted that their sites participate to attract and retain qualified staff (38 percent). A similar number joined to make the program more attractive to parents (37 percent). “We wanted the rating,” explained one center director. “We wanted to be known as a preschool that families want to come to and leave their kids.” A family child care provider added, “When parents come to see me, this gives them a little extra insight into the quality of the home and my program. More than if they just met me and thought I seem ok. I am under a microscope.”

One third of directors surveyed indicated they decided to participate to get grants and other financial incentives that the QRIS offers (35 percent). One director stated when interviewed, “[It]

[23] Appendix 6B includes several additional exhibits presenting survey response results for questions not otherwise reported in this chapter.

sounds superficial, but we ask our staff to do a lot of things and usually it's that 'here's another thing to do and then come get your same paycheck'. The idea of participating and having that incentive – it helps make staff feel like, 'OK, we can do this. It's a little over and above what we normally do, and we get rewarded for it.'" A similar percentage indicated that they chose to participate to gain more professional recognition (32 percent) or because the site was expected to participate (31 percent). See appendix exhibit 6B.5 for further details.

Financial Incentives

Fewer than half of directors reported their sites received some financial benefits for participating in the RTT-ELC QRIS, though several directors reported that they did not know if the sites they supervised had received benefits.

Directors were asked about receipt of any form of financial benefits, such as grants or awards from their site's participation in the RTT-ELC QRIS. Fewer than half of directors (42 percent) reported that their sites had received some financial benefits since enrolling in the RTT-ELC QRIS. Almost one third (31 percent; all center-based) indicated they did not know if the sites they supervised had received benefits. This response suggests that some center directors may not have complete information about the financial incentives received by the site as a result of their QRIS participation or that administrators at a higher level keep track of that information. Among directors indicating any site receipt of financial benefits, the reported use of these funds in the most recent fiscal year, July 2014 through June 2015, covered a variety of activities. Materials or curriculum purchases (67 percent) and staff training or coaching (43 percent) were the most common uses. About one quarter of sites receiving benefits reported using funds for staff bonuses or stipends or for facilities improvement (24 percent each). Within that particular fiscal year, 17 percent of site directors indicated that no benefits were received. Appendix exhibit 6B.6 provides the response rates for the full list of activities included in the survey.

To get a sense of the size of QRIS site-level financial incentives, we also gathered cost data from several Consortia for fiscal year 2014–15. That is, we asked Consortia to report the dollar amount provided through the QRIS to each site in the form of site-level financial grants or awards. Of the 11 Consortia, seven provided site-level financial incentive information, and 3 of those seven Consortia indicated that they provided $0 in site-level financial grants or awards. Among the four Consortia providing detailed data, we examined the average dollar amount per site conditional on their receipt of any dollars. The range of average site-level awards reported by Consortia for QRIS sites was $3,130 to $25,848. The mean QRIS-specific award across all sites in the four Consortia was $6,271. We note that there may be differences, however, in what award programs the Consortia included within their reported award categories—that is, it is not clear whether some staff-level awards, such as stipends for degree-level or other professional development accomplishments, may be included, which would affect the average site-level award.

Family or Public Awareness Efforts

Respondents indicated that the sites they supervised were involved in a range of parent and community engagement activities, although no single type of activity was reported by the majority.

Respondents were asked to report on the efforts that the program sites they supervised had made to inform parents and the community about their RTT-ELC QRIS involvement. The survey included a list of six parent and community engagement activities to choose from, and although respondents indicated that program sites were involved in a variety of those activities, none were reported by a majority of sites. The top two most reported efforts were each indicated by almost half of the sites: mentioning the site's QRIS involvement to new families when they enroll or inquire about the program (47 percent) and mentioning QRIS participation at parent or family engagement events at the site (46 percent). One in three sites (30 percent) mention their site's QRIS involvement in newsletters sent to parents and families. "We do talk about [quality efforts] at our parent meetings," one center director noted in an interview, "because we'll have activities that relate to say the DRDP [Desired Results Developmental Profile], [and] to help them understand and talk about how our staff is constantly getting and attending training as well…We're giving them a broader picture of the program." "I do [tell parents about quality efforts]," a family child care provider reported. "If I put it out there in my handbook and in my other paperwork that this is what we're providing here, which is what I've done, I think it lets them know that the program is improving, I am improving, so things are improving for their child and their care here."

More than one fifth (22 percent) of sites indicated that they do not currently engage in any QRIS awareness efforts. In interviews, some directors expressed frustration with parents' lack of interest in the efforts. "My parents, unfortunately, all they want to know is that the child is safe and you're doing what you're supposed to do," one family child care provider described. Appendix exhibit 6B.5 provides the response rates for the full list of awareness efforts included in the survey.

Perceptions of Tier Advancement

Directors reported that most sites without a current rating of Tier 5 were actively taking steps to move up to the next tier level. However, insufficient funding and staff education levels were reported as barriers to such tier advancement.

Among sites that did not report currently having a rating of Tier 5, 80 percent of directors responded that the sites are actively taking steps to prepare to move up to the next tier level. Only 1 percent indicated taking no active steps, and 19 percent of site respondents did not know the site's activity status at the time of the survey. See appendix exhibit 6B.7 for further details.

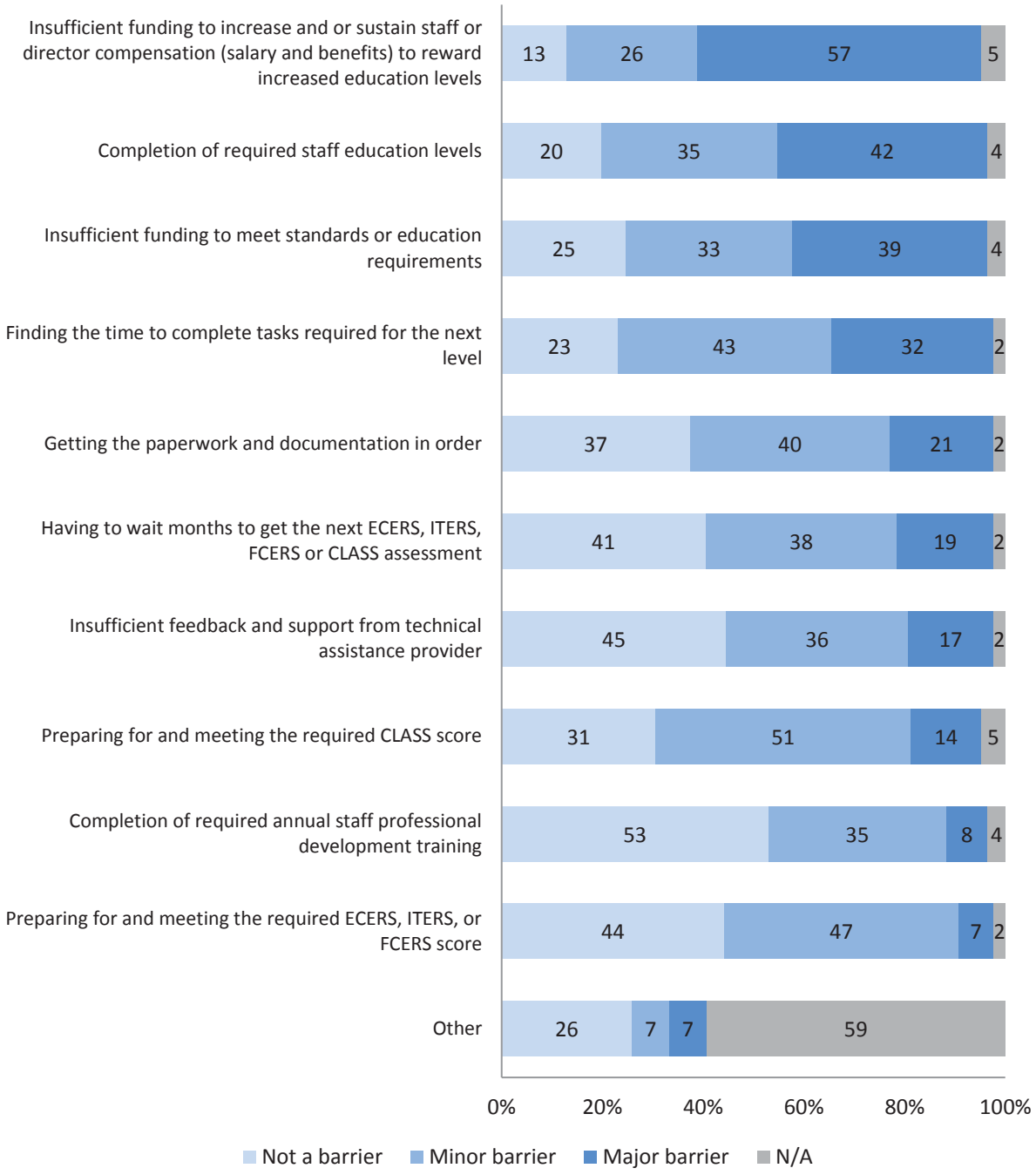
Respondents were asked to indicate which of a list of 10 barriers they perceived as potentially impeding the sites they supervised from moving up to the next tier level. Issues related to staff education levels and funding rose to the top. Exhibit 6.1 presents results for reported barriers. The most often reported (57 percent) major barrier was insufficient funding to increase or sustain staff or director compensation to reward increased education levels. "I'll hire teachers with the

education, but not the experience, so it [doesn't take long] for them to become great teachers by giving them the experience," one center director said in an interview. "But what happens is that our pay rates aren't high enough, so after a couple of years, when they get some experience behind them, they'll go somewhere that will pay them a couple more dollars. It's unfortunate but that's the way it works here." About four in 10 sites also reported that completion of required staff education levels (42 percent) and insufficient funding to meet standards or education requirements (39 percent) were major barriers.

At the same time, more than half (53 percent) of sites reported that completion of required annual staff professional development training was not a barrier. "I think [21 hours of professional development] is the right amount," one center director stated in an interview. "Our teachers do way more than that easily, but it is aligned with the 105 hours every five years for the permit process. So I like that."

We also asked directors to indicate whether each of seven RTT-ELC QRIS rating elements was especially easy or difficult to attain (FCCH respondents were not asked about two of them: ratios and group size and director qualifications) (exhibit 6.2). More site directors indicated that elements were especially easy as opposed to especially difficult. The elements reported as especially easy to attain by half or more of site directors were child observation (64 percent), ratios and group size (55 percent of centers only), and developmental and health screenings (50 percent). At least a third of sites reported that the remaining four elements were especially easy, although 15 percent of sites reported they were uncertain of their response to this question. Among those elements rated as especially difficult, the highest percentage of sites chose effective teacher-child interactions: Classroom Assessment Scoring System (CLASS) assessments (32 percent). However, a similar percentage (37 percent) indicated this element was especially easy. When asked about what prevents the site from moving up the tiers, one family child care provider interviewed reported, "The [CLASS] tool. When they come out to rate me, it's like you forgot to model this language or just doing it all the time [is hard]." "I think the big piece we're working on is teacher and child interaction," a center director reported. "The ways teachers communicate and the way they teach and what their focus is regarding the curriculum. Those are the biggest things."

Exhibit 6.1. Barriers to Moving Up to Next Tier Level: All Sites



SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.8.

NOTE: Sample includes 90 center and FCCH sites that did not report a current Tier 5 rating. Percentages are calculated based on nonmissing cases. The "Other" category had 63 missing responses, and examples of responses for "Other" include "Staff recruitment" and "Compensation for quality staff and time for planning." Responses are ordered from highest to lowest percentage by reported major barriers.

Exhibit 6.2. RTT-ELC QRIS Rating Elements That Are Especially Easy or Difficult to Attain: All

Rating Element	Easy	Difficult
Centers only (*N*=89)		
Ratios and group size	55.3	19.8
Director qualifications	37.7	18.5
All sites (*N*=102)		
Child observation	64.3	21.5
Developmental and health screenings	50.0	22.6
Minimum qualifications for lead teacher/family child care home (FCCH)	37.8	16.1
Effective teacher-child interactions: CLASS assessments	36.7	32.3
Program environment rating scale (ECERS-R, ITERS-R, FCCERS-R)	43.9	17.2
Don't know/uncertain	15.3	25.8

SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.9.

NOTE: The first two elements listed apply only to centers, so FCCH respondents did not respond to those. Respondents could select more than one answer within each column.

Only 16 to 23 percent of site respondents described each of the other six elements as especially difficult. It is interesting to note that although only 16 percent of directors highlighted attaining the minimum qualifications for lead teachers as especially difficult, 42 percent reported that completion of the required staff education levels was a major barrier to moving up the tiers. This finding may reflect some lack of clarity on the specific requirements for this element. In addition, more than one quarter (26 percent) of sites reported they did not know the answer to this question regarding ease or difficulty of meeting rating elements.

Satisfaction with the rating process varied among directors interviewed. One family child care provider noted that her rating taught her something about her program's quality, "The rating was low for me, it was anywhere between a one and a three. It was surprising because I thought I had a perfect day care." Some family child care providers with higher ratings expressed dissatisfaction with the rating process, "Can't apply the same rules to centers and family child care homes, they are two very different animals," one Tier 4 family child care provider explained. Some center directors expressed frustration with the rating criteria, "[We are] Tier 3—I felt kind of bummed about the rating. It was just two things [DRDP and staffing ratios] that tripped us up and kept us from Tier 5." Another center director explained, "We have some excellent teachers that are in classrooms with 4 stars. The ECERS [Early Childhood Environment Rating Scale] and the CLASS are just a snapshot in time… they are just one day and the scores are with her for an entire year, tied to points…. We have teachers in the four stars that are just as good as the teachers who received 5 stars."

Others were more satisfied with the rating process. "We're a 4," one center director shared. "Yes, we're pushing to be 5, but a 4 is good. I think they did a really good job of going through everything." "I believe it was overall a 3," another center director reported, "I think it was fine for where we were."

Site-Level Quality Improvement Support Efforts

Use of Administrative Self-Assessments

Half of directors reported using an administrative self-assessment at the site level in the last five years.

An administrative self-assessment, as opposed to a classroom-level assessment such as CLASS, is one way to improve overall program quality. All directors were asked which, if any, administrative self-assessment tools had been used in the last five years. Although half of site directors reported an administrative self-assessment in the past five years, nearly one third (29 percent) of directors reported no self-assessment at the site level in this time period, and another one fifth (19 percent) of directors did not know if one had taken place.

A number of tools are available to program directors who wish to improve their program through self-assessment. Directors were asked to indicate whether they were familiar with two of the more widely used self-assessment tools, the Program Administration Scale (PAS) or the Business Administration Scale (BAS). About half of the sites (49 percent) had directors reporting familiarity with either the PAS or the BAS (see appendix exhibit 6B.10). Given that only half of directors know about these assessment tools, it is not surprising that just one in five sites (21 percent) reported having used the PAS or BAS self-assessment within the past five years (exhibit 6.3). A higher percentage of sites reported using the Office of Head Start monitoring protocols (37 percent of centers). Very few reported using the National Association for the Education of Young Children (NAEYC) accreditation self-study (5 percent of centers)[24]; among FCCH sites, a small percentage reported recent use of the National Association for Family Child Care accreditation self-study (15 percent of FCCHs).

[24] Note that 12 percent of centers reported currently having NAEYC accreditation. See appendix exhibit 6B.12.

Exhibit 6.3. Completion of Administrative or Site-Level Assessments Within the Last Five Years: All Sites

	All
	Percentage
Centers (*N*=89)	
Assessment using the Office of Head Start monitoring protocols	36.5
NAEYC accreditation self-study	4.7
FCCHs (*N*=13)	
National Association for Family Child Care accreditation self-study	15.4
All sites (*N*=102)	
PAS or BAS self-assessment and continuous program quality improvement action plan	21.4
Other assessment	1.0
Don't know	19.4
No, our site has not completed a site-level assessment within the last 5 years.	28.6

SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.11.

NOTE: Includes both center and FCCH sites. Percentages are calculated based on nonmissing cases.

Use of Curriculum Frameworks

Almost all site directors expressed familiarity with the *Foundations* and *Frameworks*, and most used these frameworks for guidance on instructional practice or curriculum selection.

At the same time, almost all site directors expressed familiarity with the California Department of Education's Infant/Toddler and Preschool Learning Foundations and Curriculum Frameworks. Three quarters (75 percent) reported that their sites use these frameworks to guide their instructional practice or curriculum selection. This might reflect the large percentage of programs with Title 5 funding, which are required to have familiarity with the Foundations and Frameworks. See appendix exhibit 6B.13 for further details.

Center Requirements and Supports for Staff Professional Development Activities

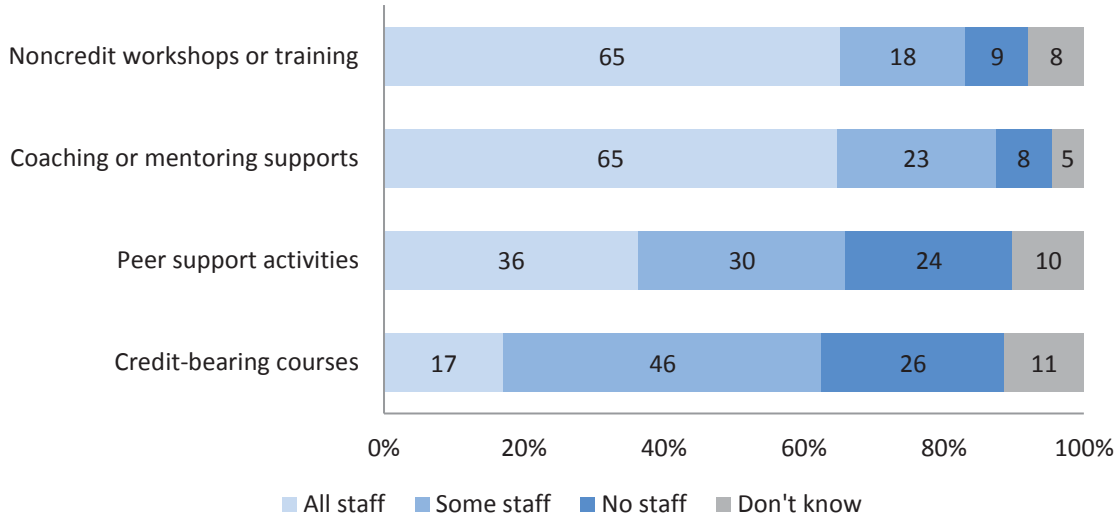
Center directors indicated that a minimum number of professional development hours is required each year for staff. Most center directors indicated that coaching or mentoring is the most helpful QI type in improving teachers' effectiveness.

We asked center directors whether their sites set requirements for annual hours of staff training, and, if so, for whom. Among centers, a minimum number of professional development hours is generally required each year for all staff types: 79 percent of centers require a minimum number of professional development hours annually for lead teachers, 63 percent for assistant teachers, and 64 percent for site administrators. See appendix exhibit 6B.14 for further details.

When asked about the centers' past year requirements for teaching staff to participate in specified professional development activities (exhibit 6.4), two thirds of center directors reported that *all* staff were required to receive coaching or mentoring (65 percent) or noncredit workshops

or training (65 percent). These percentages were lower for peer support activities (36 percent) and credit-bearing courses (17 percent).

Exhibit 6.4. Site-Required Teaching Staff Participation in Professional Development Activities by Types: Centers



SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.15.

NOTE: Percentages are calculated based on nonmissing cases. Number of missing responses varies by zero to one counts by item.

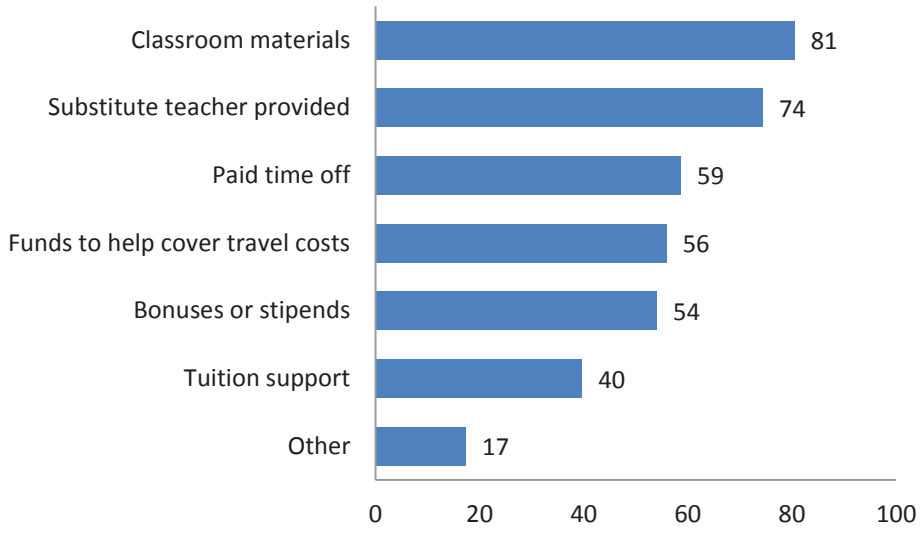
Center directors were asked for their opinion concerning which of four professional development activities is most helpful in improving a teacher's effectiveness in the classroom. Responses were overwhelmingly in favor of coaching or mentoring (74 percent). This compares with 18 percent indicating noncredit workshops or training, 6 percent indicating formal peer support, and 2 percent indicating credit-bearing courses. "It was so exciting to hear not only are you going to get this information [about quality] but you're going to have a coach that's going to guide you with this information," one interviewed center director enthused. "And then you're going to have the opportunity to work as a team ... For the staff, that was the selling point." See appendix exhibit 6B.16 for further survey details.

We also asked center directors to indicate which of six supports (e.g., paid time off, classroom materials) were offered to teachers by the site to encourage them to engage specifically in professional development or other QI activities (exhibit 6.5). Classroom materials were the most often reported support (81 percent of site respondents)[25], followed by provision of a substitute

[25] We note that for this question, we observed many missing responses for a given support type. Because of the presentation of the question, where respondents were asked to indicate yes or no for each item, it is unclear if this indicates a deliberately skipped item because the answer is no and was left unchecked, or because the respondent was uncertain of the answer. Thus, we report percentages for each item calculated based on those who answered yes or no and were not missing data.

teacher (74 percent) and paid time off (59 percent). More than half of respondents indicated that their sites offered funds to cover travel costs (56 percent) and bonuses or stipends (54 percent). Tuition support was reported by somewhat fewer respondents (40 percent).

Exhibit 6.5. Supports Offered to Encourage Participation in Quality Improvement Activities: Centers



SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.17.

NOTE: Sample includes 89 centers. Percentages are calculated based on nonmissing cases for each item.

Center Supports for Implementing New Knowledge

Center directors reported encouraging staff members to implement the knowledge gained from professional development or QI activities in a variety of ways.

Most center directors reported using a wide range of practices to encourage staff members' use of the knowledge they had gained from professional development or QI activities (exhibit 6.6). Almost all directors indicated that they encouraged their teachers to try out new ideas for their classroom (97 percent) and that they periodically observed classrooms to ensure staff are implementing new knowledge as intended (96 percent). High percentages of directors also reported checking in with staff to make sure that they have the resources needed to implement new knowledge in the classroom (89 percent) and to encourage teachers to work in teams with other staff to put new knowledge into practice (82 percent).

Almost all center directors reported that resources were made available to staff to encourage discussion and adoption of new classroom practices (exhibit 6.6). The three most commonly cited resources included classroom materials (reported for 91 percent of sites), planning time (82 percent), and teacher support staff or coaches (79 percent). Less than one third of sites reported making release time (32 percent) available to staff for the purpose of encouraging discussion or adoption of new practices.

Exhibit 6.6. Practices and Resources to Encourage Improved Classroom Instruction: Centers

Measure	Center Only
	Percentage
Which practices are used to encourage using knowledge gained from professional development or quality improvement activities?	
Encourage teachers to try out new ideas in their classrooms	96.6
Periodically observe classrooms to ensure staff are implementing new knowledge as intended	95.5
Check in with staff to make sure they have resources to implement new knowledge in the classroom	88.8
Encourage teachers to work in teams with other staff to put new knowledge into practice	82.0
Encourage teachers to discuss new ideas with their coach before implementation in the classroom	76.4
Encourage coaches/supervisors to mentor staff on how to implement new knowledge and practices	76.4
Set aside time for teachers to share knowledge with other teachers	70.8
Provide teachers planning time to turn new ideas into classroom practice	70.8
Other	6.7
What resources are available to staff to encourage discussion and adoption of new classroom practices?	
Classroom materials	91.0
Planning time	82.0
Teacher support staff, such as coaches	78.8
Teacher teams that review new practices and develop implementation plans	38.2
Release time	31.5
No specific resources like the ones listed above are available	2.3
Other	0.0
Number of sites	89

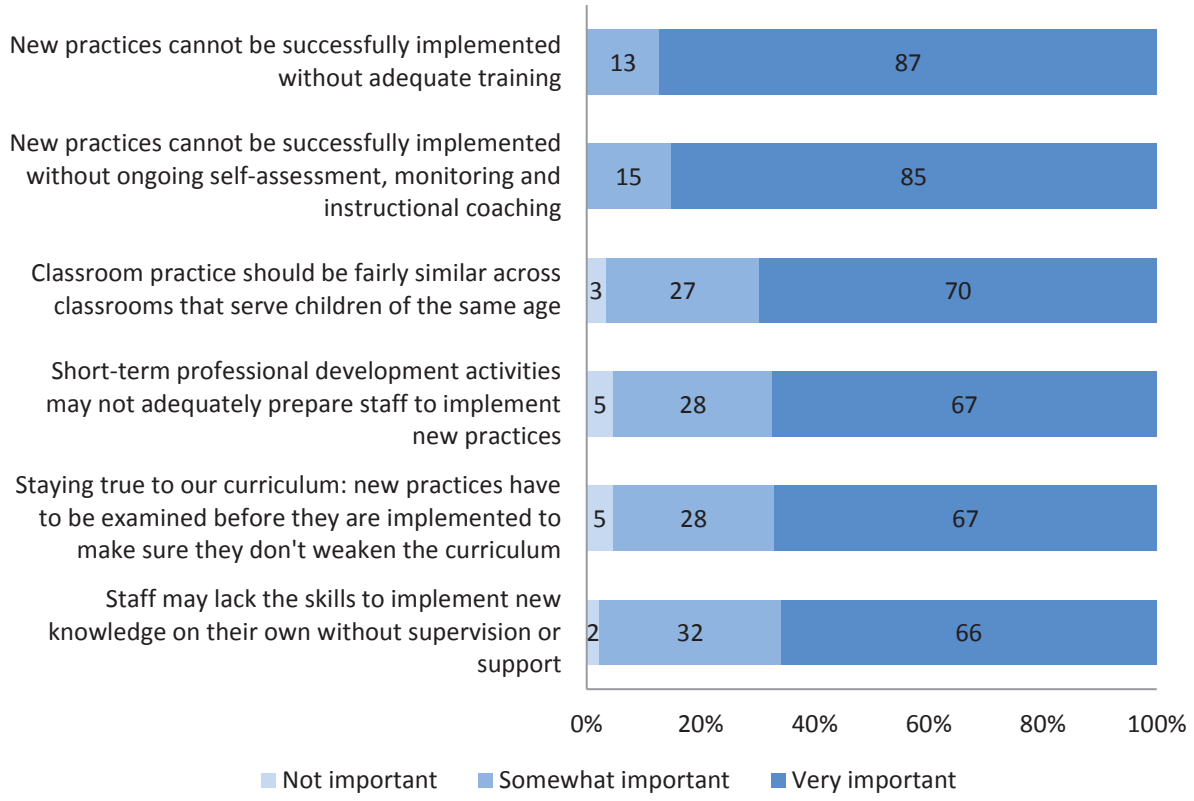
SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.18.

Though we anticipated that many site directors would see value in the idea of improving instruction by implementing new ideas in classrooms, we also recognized that several issues may need to be considered before such changes are made. We asked directors to indicate the importance of a variety of considerations—such as the need for training, ongoing monitoring, and adequacy of staff skills and training—as they thought about supporting changes to classroom instruction (exhibit 6.7). Almost all directors indicated that new practices cannot be successfully implemented without adequate training (87 percent of directors felt this consideration was very important) and without ongoing self-assessment, monitoring, and instructional coaching (reported as very important by 85 percent of directors). Indeed, very few directors considered any of the six considerations listed in the survey to be unimportant.

When asked in interviews, several directors described a community of learners at their site. For example, one center director noted, "We were looking forward to doing more training and then really going into our ECERS and ITERS [Infant/Toddler Environment Rating Scale]… really

going into it more in-depth and what it meant and how to improve the quality care that we provide. The staff was also being introduced to CLASS, so it coincided with all three of the areas we wanted to improve in the classroom." Another center director explained, "My preschool wouldn't be where it is now if it wasn't for all the help and support that we've received. It has really opened up my eyes and the teachers take advantage of the trainings and they're always looking to improve. And since we're constantly being challenged, that pushes us to make ourselves better and to be on top of whatever is the newest and latest."

Exhibit 6.7. Importance of Considerations When Supporting Changes to Classroom Instruction: Centers



SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. See appendix exhibit 6B.19.

NOTE: Sample includes 89 centers. Percentages are calculated based on nonmissing cases.

Director Quality Improvement Supports

Personal Participation in QI Activities

Almost all centers had directors who participated in personal QI activities, although completion of college courses was not as frequently reported.

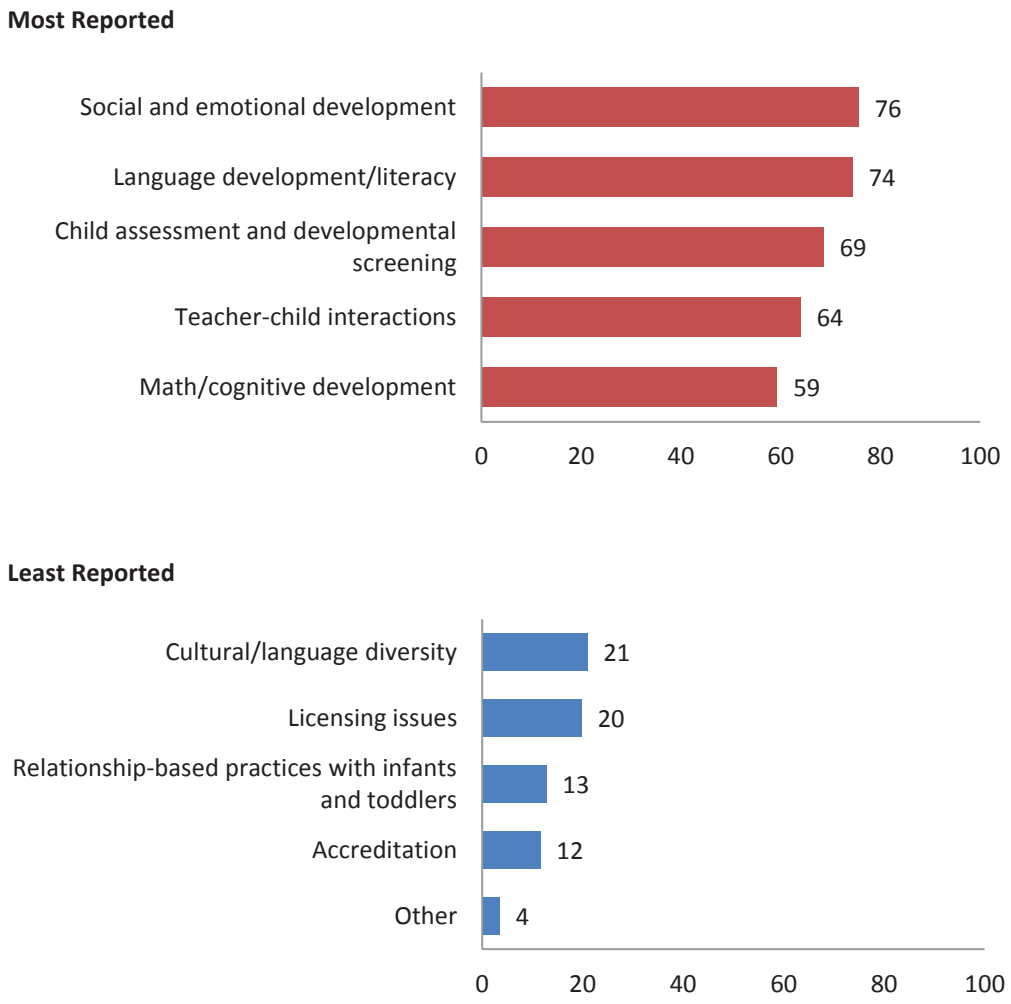
We asked site directors a set of questions about their own QI activity participation; these questions were similar to those asked in the staff survey. We report exclusively on the center

directors because FCCH respondent information on these questions was asked on the staff survey and reported in chapter 5. Most centers had directors who participated in some noncredit workshops or training (90 percent), had received some coaching or mentoring (81 percent), or had received formal peer support (80 percent) from June 2014 through March 2015. This rate of participation in coaching is very similar to what we found for staff participation rates in the staff survey, although directors' levels of participation in noncredit workshops or training and peer support are higher than those found for teaching staff. Only 20 percent of sites had directors who completed college courses during this same time frame. See appendix exhibits 6B.20 and 6B21 for further details on personal QI support.

The most frequently reported content areas covered by directors' personal QI activities were social and emotional development, language development and literacy, and child assessment and developmental screening (exhibit 6.8). About three quarters of directors reported that social and emotional development or language development and literacy were addressed by QI supports in which they had participated (76 percent and 74 percent, respectively). Many directors also reported that child assessment and developmental screening (69 percent), teacher-child interactions (64 percent), and math and cognitive development (59 percent) were addressed by their QI activities. The content areas least addressed by directors' QI support were accreditation (reported by only 12 percent of center directors), relationship-based practices with infants and toddlers (13 percent), licensing issues (20 percent), and cultural/language diversity (21 percent). For a complete list of content areas, see appendix exhibit 6B.22.

In interviews, directors highlighted the value of their own participation in the QI supports. "I wanted to see if I was on the right track myself," one center director noted. "If I was doing the right thing with the teachers and the children that we take care of... [I wanted to] try to improve in some of the areas I was lagging or not paying attention to—some of the areas I need to work on, or maybe I didn't see something that someone from the outside will be able to see it for me. They showed me what I/we need to focus on, what we had to do." Another center director stated, "I wanted to know what I was asking the staff to participate in—knowledge of all of the pieces… Also, the opportunity to know more about the CLASS, [and] those tools that may enhance our own program quality [was beneficial]."

Exhibit 6.8. Top Five Most and Least Reported QI Support Content Areas: Centers



SOURCE: Authors' analysis of the 2015 California QRIS Study Director Survey. For details, see appendix exhibit 6B.22.

NOTE: Sample includes 89 centers. Percentage of each item is calculated based on nonmissing cases.

Financial Incentives

Few center directors reported receiving any personal financial incentive for QI activity participation.

Less than one fifth (16 percent) of center directors reported receiving some form of personal financial incentive such as a scholarship or stipend for QI activity participation from July 2014 through June 2015. This is a lower rate than we found among center teacher respondents (33 percent). The value of financial incentives received was modest and similar to the average reported by center lead teachers; however, we note the small number of site directors responding to this question. Among sites with directors receiving a financial incentive in any amount, the average amount reported was $1,288 (compared with $1,328 for lead teachers), with a range of $50 to $3,300. See appendix exhibits 6B.23–6B.24 for further details.

When asked to indicate which QI efforts their financial incentives covered, almost two thirds (62 percent) of directors indicated they were used for credit-bearing courses (see appendix exhibit 6B.25). Almost 4 in 10 also reported these incentives covered coaching and noncredit workshops or training (39 percent each), and very few (8 percent) reported peer support activities.

Summary

Nearly all of the directors who responded to the survey indicated that the sites they supervise chose to participate in the RTT-ELC QRIS in order to improve program quality; substantial percentages hoped to learn new things and benefit from the technical assistance offered. Financial support was not a common reason offered for participation; in fact, fewer than half of directors reported that the sites they supervised had received financial benefits for their participation. However, this finding may not reflect the extent to which participation in the RTT-ELC QRIS overlaps with participation in other local and state initiatives that do offer financial incentives (e.g., First 5 California Child Signature Program; AB 212). It is difficult to separate the role of QRIS-specific financial incentives from the financial incentives offered by these other initiatives. Hence, it is possible that our findings regarding the role of QRIS-specific financial incentives actually understate the role of the whole cluster of financial incentives in motivating participation in QI initiatives.

The survey results suggest that programs are actively engaged in QRIS work. Four fifths of directors of sites at Tier 4 or below indicated that their sites were working to raise their tier level rating. Although many directors rated the QRIS elements as fairly easy to attain, directors cited a number of barriers to achieving higher ratings. The most frequently noted barriers were a lack of funds to raise salaries for higher educational attainment among staff and meeting standards that require higher staff education levels.

Programs also may engage in self-assessment activities as part of their QI efforts. About half of sites reported a program self-assessment within the past five years; however, nearly one third of sites reported no site self-assessment, and another one fifth of directors did not know if one had taken place. Among the several reported self-assessment tools, use of the Office of Head Start monitoring protocols was indicated most often, reflecting a common funding stream among participating sites.

Most centers set professional development standards to support their QI. Among centers, more than half had a minimum number of annual training hours required for each staff level. Two thirds of sites reported that all staff were required to receive coaching or mentoring or to participate in noncredit workshops or training. Coaching requirements are consistent with director views about its value: nearly three quarters of directors considered coaching to be the most helpful professional development activity in improving a teacher's effectiveness in the classroom. Programs also offered supports to teachers to encourage them to engage in QI activities—classroom materials was the most often reported support, followed by provision of a substitute teacher and paid time off.

Center directors also reported using a wide range of practices to encourage staff members' use of the knowledge they had gained from QI activities. Almost all directors encourage teachers to try out new ideas in their classrooms, and directors periodically observe classrooms to ensure that

staff are implementing new knowledge as intended. Nearly all directors check in with staff to make sure that they have the resources they need to implement new knowledge. Most directors reported that resources are made available to staff to encourage discussion and adoption of new classroom practices. The three most commonly cited resources include classroom materials, planning time, and teacher support staff or coaches.

Finally, most center directors also participated in their own QI activities; their level of participation in coaching was very similar to that for teachers. Less than one fifth of center directors received some form of personal financial incentive, such as a scholarship or stipend, for QI activity participation; when they did, the level was similar to amounts received by lead teachers.

These data paint a picture of active program-level QI efforts. Notable is that 80 percent of directors of programs with ratings of 4 or lower indicated that they are working to increase their rating. Moreover, despite a lack of financial incentives, most program directors reported engaging in their own QI activities. As we found with teachers, the directors' QI activities generally focus on content that is aligned with QRIS standards.

Yet, at the same time some efforts and policies such as program self-assessment and professional development standards, which would likely promote higher levels of program quality, are far from universal. State policymakers and QRIS administrators might want to consider ways to promote such efforts in programs. Also notable in our data is the relatively high percentage of directors who reported not knowing their program's tier rating. QRIS administrators may want to explore how directors learn of their ratings and how they may fail to do so. For directors and staff, a program's rating provides both critical motivation to improve and important information about where best to focus improvement efforts.

Chapter 7. Quality Improvement Activities and Changes in Teacher Practice and Children's Outcomes

Key Findings

This chapter explores the relationship between quality improvement (QI) activities and both classroom quality and child outcomes among programs participating in California's QRIS. Specifically, the classroom quality analyses examine the relationship between teacher participation in QI activities and teacher scores on the Classroom Assessment Scoring System (CLASS) instrument. The child outcome analyses examine the relationship between teacher participation in QI activities and child outcomes in the domains of preliteracy, mathematics, and executive function skills for centers only. The study examined participation in QI activities in four ways: looking at *any* participation in these activities, looking at the *amount* of participation in QI activities, looking at *sustained* participation in these activities, and looking at the topics covered in QI.

- Although we did not find consistent associations between overall participation in coaching or mentoring (whether teachers received coaching at all) and teachers' CLASS scores, the *amount* of coaching appears to matter. We found significant associations between teachers' total hours of coaching and CLASS scores as well as children's literacy, math, and executive function outcomes.

- There was some suggestion of a positive relationship between participation in peer supports and Pre-K CLASS scores, but there was no relationship with child outcomes.

- For workshops and training and ECE coursework, we found no significant relationships with CLASS scores and some negative relationships with children's outcomes, possibly reflecting the targeting of these QI supports to teachers who need the supports most.

Limitations of the study should be considered in interpreting the results. The study had only limited information about the quality improvement activities completed by programs, so the analyses cannot account for differences in the quality of training, coaching, and other quality improvement activities. Also, findings could differ if the study included a broader set of programs with more variability in their funding sources and program models.

This chapter explores the relationship between quality improvement (QI) activities and both classroom quality and child outcomes among programs participating in California's QRIS. Specifically, the classroom quality analyses examine the relationship between QI activities and teacher scores on the Classroom Assessment Scoring System (CLASS) instrument. The child outcome analyses examine the relationship between QI activities and child outcomes in the domains of preliteracy, mathematics, and executive function skills for centers only. The purpose of these analyses is to identify which QI activities are associated with improvements in teacher-child interactions in the classroom and with improvements in children's developmental

outcomes. Stakeholders in California may use this information to identify promising practices and inform closer examination of these QI activities in the future.

The analyses examine various types of QI activities that teachers may engage in, including training (noncredit workshops or training), coaching (coaching or mentoring supports), credit-bearing coursework, and peer support (see chapter 5 for definitions of each of these activities). We also examine the dosage of training, coaching, and peer support, and the content covered in training and coaching. The main analyses presented in the chapter focus on preschool teachers in centers, as they represent the large majority of teachers in our sample. However, the relationships may work differently in family child care homes (FCCHs) or among toddler teachers, although the appendices present some analyses that include FCCHs or toddler teachers. All analyses account for previous scores on the outcome measure (before the QI activities took place), as well as characteristics of the programs, teachers, and (for child outcome analyses) children. Some analyses also account for prior participation in QI activities and incentives to participate in these activities.

The reader should use caution in interpreting the findings because the study uses a design that relies on exploratory analyses that does not support the identification of causal relations between QI activities and the quality of classroom interactions or children's development. The analyses demonstrate how outcomes vary according to the ways teachers participate in QI activities, but are not designed to determine whether the QI activities themselves caused the differences in outcomes or whether other factors that affect participation in QI activities—such as teacher skills, motivation, or enthusiasm for teaching—could explain differences in outcomes. We also do not have detailed information about the quality of the QI activities such as the qualifications of the coaches and trainers or the coaching model used. It also is important to remember that the results are limited to the sample of sites participating in the study, most of which had access to QI opportunities and resources through their funding sources and the QRIS that are not available to all programs and early learning staff across the state. Results presented here should be used to inform further exploration of the relationships between QI and outcomes.

In this chapter, we address the following research questions:

- RQ 10. How do the QRIS strategies (for example, technical assistance, QI activities, incentives, compensation, and family/public awareness) improve program quality, improve the professionalization and effectiveness of the early learning workforce, and impact child outcomes? Which strategies are the least/most effective?
- RQ 11. For which QI activities does increased dosage (time and intensity of participation) impact program quality and child outcomes?

The text box below describes each of the QI activities and topics and identifies the analyses in which they are included. Additional details on the methods used in this chapter can be found in the text box and appendix 1A.

Analysis Approaches

Measuring QI activities

For this study, participating teachers completed a survey in fall 2014 and/or spring 2015, reporting information about teacher characteristics and their participation in various QI activities from July 2014 to March 2015. The analyses in this chapter include lead teachers and coteachers, but not assistant teachers. The specific QI activities included in the analyses are:

- *Participation in QI activities:* whether or not the teacher participated in peer support, training, coaching, and credit-bearing coursework (related to early care and education)
- *Dosage of QI activities:* total hours of peer support, training, and coaching received by the teacher
- *Participation in sustained coaching:* whether or not the teacher received at least 2 hours of coaching per month for at least 7 out of 10 months
- *Topics covered in training and coaching:* whether or not the teacher received a combination of both training and coaching on teacher-child interactions or understanding or improving CLASS scores (for analyses with classroom interactions), or whether or not children's language and literacy, math and cognitive development, or social emotional development were a focus of coaching (for analyses with child outcomes)

Examining the relationship between QI activities and classroom interactions:

Multiple regression analysis examined the relationship between QI activities completed by teachers and the quality of classroom interactions measured by the CLASS instrument. The analyses used cluster-robust standard errors (to account for the clustering of classrooms in centers) and full information maximum likelihood estimation (to address missing data). The analyses include CLASS score data from spring 2014 and spring 2015, as well as teacher survey data on QI activities and teacher characteristics and administrative data on program characteristics.

- Models with Pre-K CLASS scores in centers: 147 teachers (in 147 classrooms in 98 centers)
- Models with Pre-K CLASS scores in centers and FCCHs: 161 teachers (in 161 classrooms in 112 programs)

Data on QI activities and toddler CLASS scores were only available for 29 teachers, so the study team examined these data descriptively using cross-tabulations but did not conduct regression analyses with toddler teachers.

Examining the relationship between QI activities and child outcomes:

Multilevel regression models examined associations between teachers' participation in QI activities and children's skills in several domains of child development, including:

- Preliteracy skills (Woodcock-Johnson Letter-Word Identification, and Story and Print Concepts)
- Mathematics skills (Woodcock-Johnson Applied Problems)
- Executive function (Peg Tapping task)

The multilevel modeling approach accounted for the clustering of children within classrooms. Analyses drew on direct assessments of children's skills, teacher survey data, and administrative data. Analyses examined children's skills at the end of the year, controlling for children's baseline skills in the fall and other child, teacher, and program characteristics. Models included basic regression models that controlled for basic background characteristics, and additional models that also included controls for prior QI experience and receipt of financial incentives to account for teachers' access to QI, their motivation to participate, and their prior learning through those QI experiences. (See appendix 1A for more information on the child assessment instruments.)

- Models for teachers and three- to five-year-old children in centers: 1,489 to 1,552 children in 113 centers
- Models for teachers and three- to five-year-old children in centers and FCCHs: 1,547 to 1,611 children in 132 programs

Relationship Between QI Activities and Classroom Quality

We begin this chapter by examining the relationships between QI activities received by teachers and changes in their classroom quality.

Characteristics and Experiences of the Sampled Teachers

Center-based teachers in the study sample tend to be well educated and work in programs with high-quality standards, and these teachers reported a high level of participation in QI activities in the year prior to the study, so the results may not be applicable to all programs in the state.

The analyses in this section include teachers who work in a center-based program and teach in a preschool classroom or a classroom with a majority of children who are three years old or older. All teachers had a primary teaching role in their classrooms, either as a lead teacher or as a coteacher in a classroom with no one teacher designated as a lead. Exhibit 7.1 presents descriptive information about the characteristics of teachers included in the analyses in this chapter. The large majority of center-based preschool teachers, 95 percent, teach in a program that receives standards-based public funding, and more than half of them teach in a program rated a 4 or 5 in the QRIS, suggesting that most teachers are in programs with a fairly high level of quality. Sixty percent of the teachers had at least a Bachelor's degree, a relatively high percentage for early childhood teachers.

Consistent with the overall sample of teachers described in chapter 4, in 2013–14, the year prior to the study, center-based preschool teachers in this subsample reported a high level of use of most types of QI activities, particularly training (79 percent), coaching (75 percent), and peer supports (59 percent). And again, similar to the larger sample of survey respondents, fewer teachers in this subsample participated in credit-bearing coursework in 2013–14, although this is not surprising as not all early childhood teachers are in need of additional credit-based credentials, and 60 percent of the teachers had a Bachelor's degree in the 2014–15 program year.

Thus, the teachers in the study sample had access to a high level of quality supports and taught in programs meeting high-quality standards. As a result, the study analyses will describe the relationship between QI activities and CLASS scores among a sample of teachers who had access to a high level of quality supports before the study began.

Exhibit 7.2 presents teachers' current participation in center-based teacher participation in QI activities.[26]

[26] The participation rates presented in this table differ slightly from those in chapter 5 because teachers who completed a survey but did not have classroom observation data are not included here.

Exhibit 7.1. Characteristics of Center-Based Preschool Teachers in the Study Sample

Characteristic	Has Characteristic		*Total N*
	Number	Percentage	
Teacher and program characteristics in current year, 2014–15			
Teacher has at least a Bachelor's degree	82	60.29	136
English is not teacher's primary language	59	42.75	138
Teacher is White, non-Hispanic	24	17.91	134
Teaches in a program rated a 4 or 5	81	56.64	143
Teaches in a program with standards-based funding	134	95.04	141
Received a financial incentive for QI in 2014–15	49	35.51	138
QI activities in previous year, 2013–14			
Teacher participated in peer support in 2013–14	57	58.76	97
Teacher participated in training in 2013–14	96	78.69	122
Teacher participated in coaching in 2013–14	95	74.80	127
Teacher participated in credit-bearing courses in 2013–14	37	28.03	132
Characteristic	**Mean**	**Standard Deviation**	***N***
Years of early childhood education (ECE) teaching experience	9.10	10.12	136

Exhibit 7.2. Participation in QI Activities Among Center-Based Preschool Teachers in the Study Sample

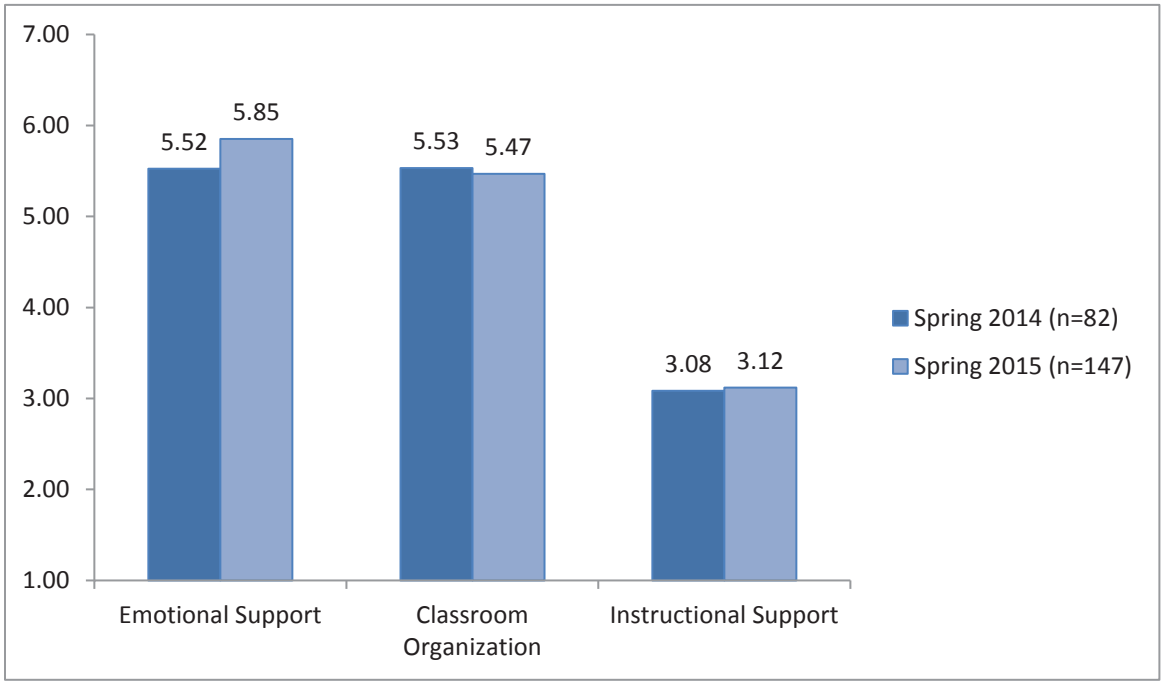
QI Activity	Participated in Activity		Total *N*
	Number	Percentage	
Participation in QI activity			
Participated in any peer supports	84	58.74	143
Participated in any training	110	76.92	143
Participated in any coaching	120	83.92	143
Participated in any credit-bearing coursework on ECE	23	17.42	132
Participation in sustained coaching			
Received at least 2 hours of coaching 7 out of 10 months	50	34.97	143
Participation in both training and coaching on topics related to classroom interactions			
Received training and coaching on teacher-child interactions or understanding or improving CLASS scores	81	60.45	134
QI Activity	**Mean**	**Standard Deviation**	***N***
Dosage of QI activity			
Hours of peer support over 10 months	11.33	17.32	139
Hours of training over 10 months	21.70	28.57	141
Hours of coaching over 10 months	20.91	32.38	143

Among teachers in the analysis sample, CLASS scores tend to be relatively high in comparison with national averages.

The study analyses examine how participation in QI activities among lead teachers in preschool classrooms relates to Pre-K CLASS scores in the three domains measured by the instrument: emotional support, classroom organization, and instructional support (see appendix 1A for a description of each domain). Scores for each CLASS domain range from 1 to 7; with 3 considered "mid-range" and 6 considered "high" quality.

Exhibit 7.3 shows that the Pre-K CLASS scores among the teachers in the study sample are relatively high in comparison with other large studies such as the Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (MS/SWEEP), which had average scores of 5.8 on the emotional support domain, 4.7 on classroom organization, and 2.1 on instructional support (Curby, Grimm, and Pianta 2010). This may in part reflect the high concentration in the sample of programs with high ratings, standards-based public funding, and a history of participation in QI activities, which may provide teachers with more support for classroom interactions and CLASS scores than do most private child care centers. The analysis results could potentially be different among a more diverse sample of teachers.

Exhibit 7.3. Average Pre-K CLASS Scores Among Teachers in the Analysis Sample, Spring 2014 and 2015



Relationships Between Types of QI and CLASS Scores

Study analyses suggest a positive relationship between participation in peer supports and teacher scores on all three domains of the Pre-K CLASS, but the relationships are not statistically significant after taking into account financial incentives and prior QI experiences.

As shown in exhibit 7.4 (and appendix exhibits 7A.1 and 7A.2), there is a statistically significant and positive relationship between participation in peer supports and scores on each of the three Pre-K CLASS domains, in analysis models that control for center teacher and program characteristics. However, these relationships remain positive but are not statistically significant after controlling for participation in peer support and other QI supports and activities in the previous year and teacher receipt of incentives to participate in QI activities. These results suggest that the positive relationship may be due in part to the peer supports themselves, but also could be due to differences between teachers who do and do not teach in programs that offer or facilitate access to peer supports. The results also suggest the possibility of a cumulative effect of peer supports over time, although additional research would be needed to examine this possibility.

Exhibit 7.4. Relationship Between Participation in QI Activities and Pre-K CLASS Scores in Centers, With and Without Controlling for Incentives and Prior QI Activities

Teacher Participated in QI Activity	Pre-K CLASS Scores, Spring 2015					
	Emotional Support		Classroom Organization		Instructional Support	
	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities
2014–15 Program Year						
Participated in peer support	0.29*	0.20	0.40*	0.11	0.65**	0.44‡
Participated in training	0.11	0.09	0.20	0.11	0.04	-0.03
Participated in coaching	-0.03	0.04	-0.33	-0.02	-0.31	-0.15
Participated in ECE courses	-0.08	-0.18	-0.21	-0.38	0.02	-0.11
2013–14 Program Year						
Participated in peer support		0.22		0.85*		0.43
Participated in training		-0.06		-0.12		-0.14
Participated in coaching		-0.04		-0.48		0.25
Participated in ECE courses		0.13		0.09		0.11

‡ $p < .10$, * $p < .05$; ** $p < .01$. $N = 147$ teachers in 98 centers.

Cells show regression coefficients, which can be interpreted as the average change in the Pre-K CLASS domain score among teachers who received each type of QI activity, compared with teachers who did not. All models control for the score on the same Pre-K CLASS domain in the previous program year (spring 2014) and for teacher and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications.

Exhibit 7.4 also shows that the analyses did not find significant differences in center-based teachers' Pre-K CLASS scores when comparing teachers who did and did not participate in training, coaching, or credit-bearing coursework on topics related to ECE. These analyses indicate whether *any* participation in each type of QI activity is related to CLASS scores. However, a limitation of these analyses is that they do not differentiate between teachers who participate in smaller versus larger amounts of professional development activities, including training, coaching, and peer support (we would not expect a large amount of variability in the amount of ECE coursework participation by actively employed teaching staff, so we are less concerned about dosage of credit-bearing coursework).

Relationships Between Dosage of QI and CLASS Scores

To better understand how the dosage of QI activities relates to subsequent CLASS scores, the study team examined the relationship between the total hours of training, coaching, and peer supports reported by center-based teachers and Pre-K CLASS scores.

There is a positive relationship between the amount of coaching reported by preschool teachers and their scores on the classroom organization domain of the CLASS.

Study analyses examining the dosage of QI activities—measured as the reported total number of hours of training, coaching, and peer supports that teachers participated in during the 2014–15 program year (from June 2014 to March 2015)—found a positive relationship between the total hours of coaching received by teachers and their scores on the classroom organization domain of the CLASS, as shown in exhibit 7.5 (and in appendix exhibits 7A.3 and 7A.4). This finding contrasts with a lack of relationship between receipt of *any* coaching and their Pre-K CLASS domain scores, suggesting that larger amounts of coaching may be needed to support change in classroom organization. Indeed, this analysis of dosage suggests that increased hours of coaching are associated with only a small increase in CLASS scores; on average, an additional 58 hours of coaching is associated with a half-point increase on the classroom organization score. The relationships between hours of coaching and scores on the other Pre-K CLASS domains are in a positive direction, but are not statistically significant.

Exhibit 7.5. Relationship Between Dosage of QI Activities and Pre-K CLASS Scores in Centers, With and Without Controlling for Incentives and Prior QI Activities

Teacher Dosage of and Participation in QI Activity	Pre-K CLASS Scores, Spring 2015					
	Emotional Support		Classroom Organization		Instructional Support	
	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities
2014–15 Program Year						
Hours of peer support	0.08	0.03	0.10	-0.02	0.06	-0.02
Hours of training	0.02	0.00	-0.11	-0.18	0.07	0.00
Hours of coaching	0.08	0.08	0.19‡	0.29*	0.12‡	0.11
2013–14 Program Year						
Participated in peer support		0.25		0.87**		0.66*
Participated in training		-0.07		-0.03		-0.08
Participated in coaching		-0.01		-0.53‡		0.21
Participated in ECE courses		0.01		-0.19		-0.13

‡ $p < .10$, * $p < .05$; ** $p < .01$. $N = 147$ teachers in 98 centers.

Cells for the 2014–15 program year show standardized regression coefficients, which can be interpreted as the average change in standard deviation units of the Pre-K CLASS domain score (one standard deviation is 0.71 for Emotional Support, 0.97 for Classroom Organization, and 1.16 for Instructional Support) for each standard deviation increase in the number of hours of training (one standard deviation is 28.57 hours of training), coaching (one standard deviation is 32.28 hours), and peer supports (one standard deviation is 17.32 hours) teachers received from June 2014 to March 2015. We report standardized coefficients here instead of regular regression coefficients because the relationship between one additional hour of each activity and the CLASS scores is small enough to round to zero in most cases, as shown in appendix exhibit 7A.3. All models control for the score on the same Pre-K CLASS domain in the previous program year (spring 2014) and for teacher and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications.

In contrast, there were no significant relationships between hours of training and Pre-K CLASS scores, or between hours of peer supports and Pre-K CLASS scores. The lack of relationship between hours of peer supports and CLASS scores is surprising as there did appear to be a small positive relationship between receipt of any peer supports and CLASS scores, suggesting that access to any peer supports may be more meaningful than the specific amount received by teachers in programs that offer peer supports.

The positive relationship between hours of coaching and the classroom organization domain of the Pre-K CLASS raises the question of how much coaching is needed to be helpful. This analysis of coaching dosage does not differentiate between teachers who received coaching during a narrow window of time and teachers who received sustained coaching throughout the program year. To better understand how participation in sustained coaching relates to subsequent CLASS scores, the study team compared Pre-K CLASS scores of teachers with and without coaching over the course of the program year.

Participation in ongoing, sustained coaching over the course of the program year was associated with higher scores on the Emotional Support domain of the Pre-K CLASS, although the relationship is only marginally significant after controlling for prior participation in coaching and other supports.

As shown in exhibit 7.6, analyses examining participation in sustained coaching—measured as receiving at least two hours of coaching per month for at least seven of the 10 months covered by the survey—find a significant and positive association with Pre-K CLASS emotional support scores. Among teachers in centers, this relationship became only marginally significant after controlling for participation in coaching in the prior program year. There was no relationship between participation in sustained coaching and scores on other Pre-K CLASS domains.

Exhibit 7.6. Relationship Between Participation in Sustained Coaching and Pre-K CLASS Domain Scores in Centers

Teacher Dosage of and Participation in QI Activity	Pre-K CLASS Scores, Spring 2015					
	Emotional Support		Classroom Organization		Instructional Support	
	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities	Controlling only for teacher character-istics	Also controlling for incentives and prior activities
2014–15 Program Year						
Received at least 2 hours of coaching per month, 7 of 10 months	0.25*	0.23‡	0.11	0.03	0.23	0.06
2013–14 Program Year						
Participated in peer support		0.28		0.96***		0.67*
Participated in training		-0.12		-0.27		-0.27
Participated in coaching		0.00		-0.42		0.28
Participated in ECE courses		-0.01		-0.15		-0.04

‡ $p < .10$, * $p < .05$, ** $p < .01$. $N = 147$

Cells show regression coefficients, which can be interpreted as the average change in the Pre-K CLASS domain score among teachers who received sustained coaching, compared with teachers who did not. All models control for the score on the same Pre-K CLASS domain in the previous program year (spring 2014) and teacher and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications.

Relationships Between the Focus of QI and CLASS Scores

The study team also examined the relationship between receipt of both training and coaching on topics related to classroom interactions and Pre-K CLASS scores, under the assumption that receiving combined training and coaching on this topic would provide particular support for QI.

There was no association between receipt of both training and coaching on topics related to classroom interactions and subsequent Pre-K CLASS scores.

The study found no significant relationships between the combined receipt of both training and coaching on topics related to classroom interactions in the 2014–15 program year and Pre-K CLASS scores in spring 2015. Specifically, the study examined the teacher's report of receiving both training and coaching on teacher-child interactions or understanding or improving CLASS scores (see appendix 7A).

Relationships Between QI Activities and CLASS Scores Among FCCHs and Toddler Teachers

Results were similar in supplemental analyses that combined teachers in FCCHs with preschool-age children with preschool teachers in centers.

The analyses presented in this chapter do not include teachers in FCCHs because there are differences between centers and FCCHs in terms of program and teacher characteristics, as well as access to QI activities, suggesting that the relationship between QI activities and CLASS scores might differ in substantive ways and should be examined separately for centers and homes. The analyses could not be performed separately for staff in homes because of the small number of study participants who teach in FCCHs and have predominantly preschool-age children, but supplemental analyses that included both centers and homes had very similar findings to those presented in this chapter (see appendix 7A).

Results may differ for toddler teachers, but the small number of toddler teachers and differences in data for these teachers did not permit us to examine this relationship empirically.

Analyses examining the relationship between QI activities and CLASS scores could not include toddler teachers because the domains measured for toddler classrooms differ, but descriptive analyses suggest that the small number of toddler teachers in our sample differed from preschool teachers in terms of program characteristics (see appendix exhibit 7A.9), and that there also were some differences in the QI activities they received (see appendix exhibits 7A.10 and 7A.11). In particular, the toddler teachers in our sample appear to be less likely to receive peer supports than the preschool teachers in our sample—although differences between toddler and preschool teachers should be interpreted with caution given the small number of toddler teachers. Still, these apparent differences suggest that the relationship between these activities and CLASS scores may work differently in toddler classrooms. However, we are not able to test this possibility empirically because of the small number of toddler teachers.

Relationship Between QI Activities and Children's Outcomes

This section of the chapter presents results of models examining associations between teacher participation in QI activities and children's developmental outcomes related to literacy, mathematics, and executive function (social-emotional skills).

Characteristics of the Sampled Teachers and Children

As in the analyses presented above, the sample consists of teachers who were in a primary teaching role (i.e., lead or coteachers) in a center-based preschool classroom with children ages three years and above.[27] These analyses focus on a subset of those teachers who completed a survey and whose children were assessed by the study team. (See appendix 7B for further details on the teacher sample.) The child sample included boys and girls in nearly equal proportions. Approximately 9 percent of the sample children had special needs. The majority of children in the sample spoke Spanish at home, either alone or in combination with English, while less than a third of the children spoke exclusively English at home. Sixty-three percent of the children in the study sample were English proficient. These children participated in assessments of preliteracy skills (Woodcock-Johnson Letter-Word Identification subtest and Story and Print Concepts), mathematics skills (Woodcock-Johnson Applied Problems subtest), and executive function (Peg Tapping task). Descriptive statistics for children's assessments in the fall and spring are presented in exhibit 7.7 to provide context for the results of regression models. Exhibit 7.8 provides an overview of the child assessment instruments. Appendix 1A provides more detailed information on the assessment instruments and procedures.

[27] Family child care providers were not included in the analyses presented in the chapter because (1) they have lower QRIS ratings than centers, on average; (2) they have a different pattern of QI participation; and (3) they are small in number in our analytic sample. Analyses that include family child care providers are included in appendix 7C for the reader's reference.

Exhibit 7.7. Characteristics of Children Enrolled in Center-based ECE Programs in the Study Sample

Characteristic	Has Characteristic		*N*
	Number	Percentage	
Child is male	545	50.70	1,075
Child has special needs	99	9.21	1,075
Home language			
English	325	30.23	1,075
Spanish	373	34.70	1,075
English and Spanish	314	29.21	1,075
Other	63	5.86	1,075
Sufficiently proficient in English to be assessed in English	676	62.88	1,075
Characteristic	**Mean**	**Standard Deviation**	***N***
Fall Scores			
preLAS (Sum of items correct score range: 0-20)	12.27	(6.21)	1,075
Peg Tapping Task (Sum of items correct score range: 0-16)	6.48	(4.92)	1,065
Story and Print Concepts (Sum of items correct score range: 0-8)	3.30	(2.14)	1,065
Letter-Word Identification (Age equivalent score range: 2-28)	4.39	(1.11)	1,061
Applied Problems (Age equivalent score range: 2-28)	3.30	(0.86)	1,046
Spring Scores			
preLAS	15.15	(4.99)	1,075
Peg Tapping Task	8.87	(5.43)	1,072
Story and Print Concepts	4.45	(2.28)	1,072
Woodcock Johnson Letter-Word Identification	4.95	(1.07)	1,066
Woodcock Johnson Applied Problems	3.76	(0.83)	1,064

Relationships Between Types of QI and Child Assessment Scores

Study analyses suggest a negative relationship between current (as opposed to prior) teacher participation in credit-bearing ECE courses and children's literacy outcomes, and mixed relationships between coaching and literacy outcomes, but findings should be considered in light of potential differences between teachers who do and do not participate in coursework intended to increase their qualifications.

Children whose teachers participated in ECE coursework in the current program year (that is, over the past 10 months prior to the 2015 spring child assessments) had lower literacy scores, on average, than children whose teachers did not. This negative association was observed with the Story and Print Concepts measure and on the Woodcock Johnson Letter-Word Identification subtest (exhibit 7.8). The negative relationships between teacher coursework in ECE and children's literacy outcomes were stronger after controlling for teachers' participation in QI activities in the previous year, although there appeared to be a positive relationship between teacher coursework in ECE in the previous year (2013–14) and child outcomes on these literacy skills in spring 2015. In addition, teacher participation in ECE coursework in 2014–15 was

negatively associated with children's development of early mathematics skills, although this negative association was no longer statistically significant after controlling for teachers' previous participation in QI activities. The negative association between ECE coursework and child outcomes is not inconsistent with analyses examining CLASS score outcomes; although there was no statistically significant relationship between ECE coursework in 2014–15 and CLASS scores, the direction of the relationships appeared to be negative.

Exhibit 7.8. Associations Between Teachers' Participation in QI Activities and Child Outcomes, With and Without Controls for Incentives and Prior QI Activities

	Child Outcomes, Spring 2015							
	Story and Print Concepts		Peg Tapping Task		Letter-Word Identification		Applied Problems	
	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI
2014-15 Program Year								
Training	-0.130	0.072	-0.266	0.037	-0.122‡	-0.026	0.014	0.088
Coaching	-0.348	-0.462*	-0.038	-0.443	0.241**	0.164*	0.058	0.011
ECE courses	-0.885***	-0.974***	0.231	0.015	-0.120	-0.172*	-0.117*	-0.085
Peer support	0.160	0.103	0.469	0.390	-0.023	-0.112	0.010	-0.031
2013–14 Program Year								
Training		-0.471		-0.169		-0.353**		-0.225**
Coaching		0.569*		1.422**		0.370**		0.062
ECE courses		0.530***		0.079		0.287**		0.056
Peer support		-0.177		-0.971*		0.044		0.141‡

‡ $p < .10$, * $p < .05$; ** $p < .01$. N = 1,037 to 1,064 children, taught by 108 teachers, in 87 centers.

Cells show regression coefficients, which can be interpreted as the average change in child outcome scores in spring 2015 among children with teachers who participated in each type of QI activity, compared with children with teachers who did not. All models control for child scores on the same assessment in fall 2014 and child, teacher, and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications. See appendix 7B for the full set of regression coefficients.

The study design does not allow us to determine why current coursework is negatively related to literacy outcomes while prior coursework is positively related. It is possible that the skills teachers learn from ECE coursework are delayed (for example, it could take a year or more for the skills learned in ECE coursework to be implemented successfully in working with children). Teachers who are juggling a full time job and taking classes in the evening might also find it difficult to give their full attention to their classroom during the day. It may also be that there are differences between teachers who are participating in coursework to increase their qualifications as early childhood teachers compared to those who are not. As with all analyses in this study, we do not examine cause-and-effect relationships because we are unable to account for all meaningful differences between teachers who did and did not participate in ECE coursework. Although the study analyses control for teachers' years of experience and degree level, there may be other differences between teachers who are and are not participating in this type of

coursework—or between children who are and are not in these teachers' classrooms—that could explain differences in child outcomes between the two groups.

These analyses had mixed findings regarding the relationship between participation in coaching and child literacy outcomes. The results suggest that teacher participation in any coaching during the current program year and in the prior program year was positively associated with children's development of letter and word identification skills. However, participation in any coaching in 2014–15 was negatively associated with children's familiarity with story and print concepts after controlling for prior participation in QI activities, but prior participation in coaching was positively related to this outcome. As with ECE coursework, the study design does not allow us to determine whether these differences are explained by differences in the teachers or children being compared, or whether there is an alternative explanation.

There was no observed association between teacher participation in peer support and any child outcomes. Study analyses examining Pre-K CLASS score outcomes found a positive association between participation in peer supports and CLASS scores across domains, but the relationships were smaller and not statistically significant after controlling for prior participation in QI, so these results are consistent. Also, there was no association between teacher participation in training and any child outcomes, consistent with the observed lack of association with CLASS score outcomes.

Relationships Between Dosage of QI and Child Assessment Scores

Participating in more hours of coaching in the current year is consistently positively associated with children's literacy, mathematics, and executive function skills, even after controlling for prior participation in QI activities.

In addition to any participation in QI activities, the study team also examined how the dosage of QI (defined as the total number of hours over the previous 10 months of coaching, training, and peer supports) relates to child developmental outcomes. Analyses suggest that the total number of hours of coaching that teachers received during the current program year showed a small positive association with children's executive function, ability to identify letters and words, and early mathematics skills, as shown in exhibit 7.9.[28] The positive association between total hours of coaching and letter-word identification is consistent with the positive relationship observed for any participation in coaching; in contrast, the negative association between participation in any coaching and story and print concepts is not observed when examining the relationship between hours of coaching and this outcome. The positive relationship between hours of coaching and the executive function and early mathematics skills outcomes contrasts with the lack of association for participation in any coaching, suggesting that the amount of coaching matters for these outcomes. The largely positive relationship between hours of coaching and child outcomes is consistent with a positive association between hours of coaching and classroom organization scores on the Pre-K CLASS. However, selection bias may be influencing our findings, since

[28] The coefficients are numerically small because they examine the relationship between each hour of QI support and child outcomes. For example, 10 additional hours of coaching would increase scores on the Peg Tapping task by 0.18.

teachers who participate in more hours of coaching may be fundamentally different from those who receive fewer hours of coaching. They may be more motivated, more persistent, or more skilled in working with children.

Taking into account prior QI activities, participating in more hours of workshops or training is negatively associated with children's executive function.

Although there was no relationship between any participation in training and child outcomes, there were observed negative associations between the total number of hours of training that teachers received in the current program year and children's executive function skills, after controlling for prior participation in QI (see exhibit 7.9). In addition, we observe a negative association between hours of training and both early literacy outcomes, which was not statistically significant after controlling for prior participation in QI. However, for most child outcomes, there was a negative relationship for participation in training in the prior program year. As discussed previously, these analyses do not identify cause-and-effect relationships and could be explained by differences in teachers or children being compared. For example, one possible explanation for this finding is that teachers who work with students who have lower levels of executive function may attend more hours of training.

Exhibit 7.9. Associations Between the Number of Hours of QI Activities Teachers Receive and Child Outcomes, With and Without Controls for Incentives and Prior QI Activities

	Child Outcomes, Spring 2015							
	Story and Print Concepts		Peg Tapping Task		Letter-Word Identification		Applied Problems	
	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI
2014–15 Program Year								
Hours of training	-0.006**	-0.002	-0.007‡	-0.010*	-0.002**	0.000	0.000	0.001
Hours of coaching	0.003	0.002	0.020**	0.018**	0.004**	0.002**	0.003**	0.002**
Hours of peer support	0.000	-0.001	-0.009	-0.009	-0.001	-0.002	-0.001	-0.001
2013–14 Program Year								
Training		-0.814**		-0.044		-0.293*		-0.214*
Coaching		0.420		0.973‡		0.259‡		-0.006
ECE courses		0.257		0.037		0.212**		0.013
Peer support		0.144		-0.613		0.021		0.165*

‡ $p < .10$, * $p < .05$; ** $p < .01$. N = 1,037 to 1,064 children, taught by 108 teachers, in 87 centers.

Cells show regression coefficients, which can be interpreted as the average change in child outcome scores in spring 2015 for each hour that teachers participated in each type of QI activity. All models control for child scores on the same assessment in fall 2014 and child, teacher, and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications. See appendix 7B for the full set of regression coefficients.

Participation in sustained coaching is positively associated with children's development in most domains.

Children whose teachers received sustained coaching (for two or more hours a month for at least seven months over the course of the current program year) had higher average skills in the domains of executive function, letter and word identification, and early mathematics, relative to children whose teachers did not receive sustained coaching (see exhibit 7.10). Results for executive function and letter and word identification persisted after controlling for previous participation in QI activities. These results are largely consistent with the positive relationships observed between the total hours of coaching and child outcomes, although the mathematics outcome is not statistically significant after controlling for prior QI participation. The positive associations with child outcomes is consistent with the positive direction of the relationship between sustained coaching and Pre-K CLASS scores, although the CLASS outcomes were not significant in any domain after controlling for prior participation in QI. However, we must keep in mind the possibility of selection. Teachers who participate in sustained coaching over the course of the program year may be more motivated or effective than those who do not.

Exhibit 7.10. Associations Between Sustained Coaching and Child Outcomes, With and Without Controls for Incentives and Prior QI Activities

	Child Outcomes, Spring 2015							
	Story and Print Concepts		Peg Tapping Task		Letter-Word Identification		Applied Problems	
	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI
2014–15 Program Year								
Sustained coaching	-0.167	-0.321‡	0.579*	0.544‡	0.278**	0.215**	0.111**	0.066
2013–14 Program Year								
Training		-0.937**		-0.270		-0.271‡		-0.208*
Coaching		0.552*		1.121*		0.189		0.001
ECE courses		0.360*		0.134		0.203**		0.015
Peer support		0.195		-0.715‡		0.007		0.166*

‡ $p < .10$, * $p < .05$; ** $p < .01$. N = 1,037 to 1,064 children, taught by 108 teachers, in 87 centers.

Cells show regression coefficients, which can be interpreted as the average change in child outcome scores in spring 2015 among children with teachers who participated in sustained coaching, compared with children with teachers who did not. All models control for child scores on the same assessment in fall 2014 and child, teacher, and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications. See appendix 7B for the full set of regression coefficients.

Relationships Between the Focus of QI and Child Assessment Scores

Coaching focused specifically on language and literacy was positively associated with children's literacy skills; however, participation in coaching specifically focused on mathematics or cognitive development was negatively associated with children's mathematics skills.

To take a more nuanced look at coaching, the study team also examined associations between coaching focused on specific content areas and outcomes in those same content areas. Results indicated that children did better on letter and word identification when their teachers received coaching focused on language and literacy (at least 25 percent of their coaching time focused on language and literacy), although there is no association between focused coaching on this topic and child outcomes on the story and print concepts measure (see exhibit 7.11.). Conversely, children scored lower on the applied problems measure of mathematics skills if their teacher received coaching focused on mathematics or cognitive development (at least 25 percent of their coaching time on this topic). There was no relationship between focused coaching on social-emotional development and children's executive function.

Exhibit 7.11. Associations Between Focused Coaching on Specific Topics and Child Outcomes, With and Without Controls for Incentives and Prior QI Activities

	Child Outcomes, Spring 2015							
	Story and Print Concepts		Peg Tapping Task		Letter-Word Identification		Applied Problems	
	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI	Basic Model	With Incentives and Prior QI
2014–15 Program Year								
Focus of coaching (25% or more time)								
Language/Literacy	0.095	0.108	-0.098	0.143	0.173	0.189*	-0.004	0.030
Math/Cog. dev.	-0.088	-0.261	0.225	-0.348	0.058	-0.041	-0.008	-0.085*
Soc. emo. dev.	-0.212	-0.232	-0.301	-0.527	0.013	-0.009	0.040	-0.014
2013–14 Program Year								
Training		-0.924**		-0.745		-0.349*		-0.274**
Coaching		0.514‡		1.708**		0.301*		0.084
ECE courses		0.283‡		0.104		0.224**		0.021
Peer support		0.234		-0.311		0.038		0.204**

‡ $p < .10$, * $p < .05$; ** $p < .01$. N = 1,037 to 1,064 children, taught by 108 teachers, in 87 centers.

Cells show regression coefficients, which can be interpreted as the average change in child outcome scores in spring 2015 among children with teachers who participated in coaching focused on select topics, compared with children with teachers who did not. All models control for child scores on the same assessment in fall 2014 and child, teacher, and program characteristics. Models indicated as controlling for prior QI activities also control for participation in training, coaching, credit-bearing ECE courses, and peer support in the prior program year, 2013–14, as well as receipt of a financial incentive for QI activities in the 2014–15 program year. See appendix 1A for additional detail about model specifications. See appendix 7B for the full set of regression coefficients.

The study design does not allow us to determine why more intensive coaching on language and literacy would be associated with better early literacy outcomes for children, but more intensive coaching on mathematics and cognitive development would be associated with poorer child outcomes for mathematics. As discussed previously, the study analyses do not examine cause-and-effect relationships, so the relationships could be explained by differences between teachers who did and did not receive coaching on each topic, or the children they teach. For example, coaches may focus on mathematics instruction with the teachers who most struggle to teach this content. Descriptive statistics for the analytic sample indicate that teachers whose coaching included a greater focus on mathematics were less likely to hold a Bachelor's degree, which may indicate a weaker background in mathematics and/or mathematics education overall. Thus, these may be the teachers who need more support teaching mathematics content.

Summary

In summary, the study examined the relationship between participation in and the amount of QI activities and both classroom quality and child developmental outcomes.

For coaching, the study team found positive associations between the total hours of coaching teachers received and child outcomes, including letter-word identification, executive function, and early mathematics skills. We also found positive association between participation in sustained coaching over the year and letter-word identification and executive function. In contrast, we found that participating in any amount of coaching was associated with letter-word identification, but was not associated with executive function or mathematics skills. This suggests that for these skills in particular, it may be that small amounts of coaching are not adequate to support teachers in helping children learn in these more challenging areas of child development. However, the analyses do not examine cause-and-effect relationships, so more research is needed on whether larger amounts of coaching or sustained coaching cause improvements in child outcomes or whether some other explanation such as teacher motivation is related to both coaching participation and child outcomes.

We also found a positive association between the total hours of coaching and the classroom organization domain of the Pre-K CLASS instrument, which measures the quality of classroom materials available to children; the time children spend engaged in tasks; and the rules, expectations, and behavior management. The study found some suggestion of a positive relationship between sustained coaching over the year and the emotional support domain of the CLASS, which measures teacher sensitivity and the classroom climate, but the relationship was not statistically significant after controlling for prior participation in QI activities and receipt of incentives. In contrast, analyses found no significant differences in CLASS scores for teachers who received any coaching versus those who received no coaching, again suggesting that small amounts of coaching may not be enough to support improvements in classroom quality. Additional research is needed to examine whether this is the case.

In addition to positive relationships between coaching in the current year and several study outcomes, the analyses also found some positive relationships between participation in coaching in the previous year and child outcomes.

Although the study largely found positive relationships between coaching and child outcomes, the outcome related to story and print concepts was an exception. Study analyses found a negative association between participation in any coaching and child scores on this measure, and found nonsignificant relationships that appeared to be negative in direction for both hours of coaching and sustained coaching. These inconsistent relationships are difficult to understand but may be an indication that these book knowledge skills (for example, identifying the author and title of the book) were deemphasized in coaching sessions.

For the topics covered in QI activities, the study found that intensive coaching on language and literacy is associated with better early literacy outcomes for children on the letter-word identification measure, but also found that intensive coaching on mathematics and cognitive development is associated with worse child outcomes for mathematics. The study design does not allow us to determine the reason for these relationships, but possible explanations could include differences between teachers who do and do not receive coaching on specific topics or between the children they teach, or possibly that coaches may be better prepared to support teachers in literacy than in mathematics. Additional research is needed to determine how the topics of coaching affect child learning and developmental outcomes. There were no differences in CLASS scores between teachers who did and did not participate in both training and coaching focused on classroom interactions or understanding or improving CLASS scores.

For peer supports, the study found no relationship with child outcomes. There was an observed positive relationship between any participation in peer supports and all Pre-K CLASS domains; however, that became smaller and nonsignificant after controlling for prior participation in QI activities and receipt of incentives. There was no observed relationship between the number of hours of peer supports and CLASS or child outcomes. In summary, the study found no clear indication of a benefit to peer supports, but this cannot be determined with certainty as the study is not designed to examine cause-and-effect relationships.

For participation in ECE coursework, the study found no significant relationship between participation in the current year and CLASS scores, and a significant negative relationship with literacy outcomes. However, participation in ECE coursework in the prior program year has a positive relationship with these same literacy outcomes. The study design does not allow us to determine if this finding is due to differences between those who did and did not participate in coursework, or if there is an alternative explanation for this counterintuitive finding, such as a time lag between the coursework itself and implementation of what was learned to support child learning. We suggest further research to examine how and when participation in ECE coursework affects child learning.

For training, the study analyses found that participation in any training in the current program year was not related to child outcomes. However, the total hours of training in the current year was negatively associated with executive function and appears to have a negative but nonsignificant relationship with literacy outcomes. Furthermore, participation in training in the prior program year was negatively associated with child outcomes in most models. The study found no relationship between training and any CLASS outcomes. It is important to note that we do not have information about the quality of the training or the extent to which training experiences were provided with fidelity to the program model and intended goals.

In all analyses in this study, we cannot identify cause-and-effect relationships because we are unable to account for all meaningful differences between teachers who did and did not participate in different types or amounts of QI—or between children who attend programs in which the teachers receive differing amounts of QI. Although the study analyses control for some teacher and child characteristics, there may be other differences between teachers or children that could explain the observed relationships. There is likely significant variation in the content and quality of QI experiences, in teachers' skills and needs, and in children's learning needs, and the lack of associations between some QI activities and outcomes could reflect a misalignment of these factors. These analyses provide initial information about how QI activities are related to child outcomes, and further research is needed.

Also, the teachers participating in the study had a relatively high level of education, and most taught in high-rated programs with standards-based public funding. Thus, the relationships observed in this sample may not apply to all early childhood teachers in the state or to all children in early childhood education programs in the state.

Chapter 8. Cost of QRIS Consortia Quality Improvement Supports

Key Findings

This chapter provides information on the economic costs associated with the major quality improvement (QI) strategies examined in the study: coaching or mentoring, credit-bearing courses, noncredit workshops or training, peer support activities, and financial incentives. We examined the per-unit cost of various QI activities, using data from a cost survey collected from 11 Consortia.

- Among the few Consortia with the most reliable data, the cost per participant averaged $3,400 for coaching, $1,350 for credit-bearing courses, just over $300 for noncredit workshops or trainings, and just over $2,400 for peer supports.

- Thus, coaching, the type of QI activity that was most consistently related to both quality outcomes (that is, CLASS scores) and children's developmental outcomes, also was the most expensive. This is not surprising, given that, unlike the other types of QI activities, coaching is typically conducted one-on-one and thus is more labor intensive than group-administered workshops, which did not show positive associations with quality or children's developmental outcomes.

Given the limitations on the cost information provided by the Consortia, the estimated cost figures should be viewed as approximations of the true economic cost. Furthermore, there is considerable variation across Consortia in the resulting cost estimates, perhaps due to variation in approaches to QI activities, but some of the variation may also reflect the challenge of calculating such costs in a consistent manner.

This chapter provides information on the economic costs associated with the major quality improvement (QI) strategies examined in the study: coaching or mentoring, credit-bearing courses, noncredit workshops or training, peer support activities, and financial incentives (see side bar) and addresses components of the following research question:

- RQ 13. What is the cost versus benefit for various QRIS strategies relative to child outcomes?

In administering the QI Activity Cost Survey developed for this project (see appendix 8 for more details on the survey), the goal was to collect information from Consortia using a standardized protocol that would allow us to estimate the cost per unit of QI activity or output (for example, cost per coaching hour, cost per credit-bearing course) across the 11 focal Consortia.

To the extent possible, the administrator in each local Consortium was asked to provide a comprehensive accounting of the annual expenditures to implement QI supports—including cash outlays, as well as in-kind contributions—and to allocate those costs across each type of QI support provided by the local Consortium. In addition, we requested that each Consortium report on the level of output associated with each type of QI support (for example, the number of program staff receiving coaching, the total number of coaching hours provided, and so on). Although the QI support cost and output information was collected for each Consortium, results are presented in summary form (for example, the average and range). We did not seek to explain the variation in cost per unit of QI activity that we estimated because that was beyond the scope of our analysis.

To our knowledge, this is one of the first efforts to quantify the economic cost of implementing QI activities in the context of local Consortia. Generating cost estimates for QI activities can be of interest in their own right for understanding the resource requirements to support QI within the context of a local Consortium. In addition, when estimates are available of the impacts of QI supports on program quality improvement or children's developmental gains associated with each QI activity, information on the cost of those QI activities can be used to compare the relative cost-effectiveness of each type of QI support.

Analysis Approaches

- Descriptive analysis of cost data collected from focal Consortia to estimate the cost per unit of QI activity.

Data and Sample

- Data collected from 11 focal Consortia on the cost of QI activities and associated QI activity output.

Definitions of QI Activities

- ***Coaching/mentoring***: Programs that provide support for individualized professional development, usually one-on-one or as part of a classroom team, provided by a coach, mentor, or advisor, to help improve early educators' practice or to promote QI more generally.
- ***Credit-bearing courses***: Courses completed for unit credit at a two- or four-year college or university.
- ***Financial incentives***: Financial incentives to promote participation in QI efforts or to improve the program more generally. This category can include wage supplements, tiered reimbursement, quality awards, and so on. Financial incentives represent those not directly attributable to another QI strategy.
- ***Noncredit-bearing trainings***: Courses, seminars, workshops, or training programs that may be one-time or part of a series (courses that provide continuing education units but not those taken for formal college credit through a college or university).
- ***Peer support activities***: Formal arrangements such as learning communities, peer support networks, or reciprocal peer coaching to discuss shared experiences and exchange ideas, information, and strategies for professional development or for program improvement more generally.

Given that most of the local Consortia we examined conduct multiple QI activities, it often was challenging to allocate the shared cost of administration and other overhead activities across each of the QI components. In addition, the QI supports in each local Consortium typically involved some combination of state-funded and locally funded activities, where the state-funded activities may be administered at the state level or the local level. This made it difficult to ensure alignment between the expenditure amounts reported for any given QI support on the part of the local Consortia and the level of QI outputs, which may include activities supported with state funds.

It should be noted that the survey gave Consortia discretion in categorizing costs by QI activity. Definitions were provided for each QI support to guide Consortia administrators in their classification of programs and activities (see sidebar). For each Consortium, we included a table of known programs and how, based on our understanding of the program, each program should be grouped, but Consortia administrators were ultimately responsible for classifying the programs/activities included in the cost totals. We provided feedback and answered questions if Consortia were uncertain how costs should be classified.

In the results presented in the next section, we use the reported data collected in the survey to generate estimates of the costs per unit of QI activity as expended through Consortia funds. We primarily focus on reported figures from five Consortia with the most comprehensive and reliable information. As described more fully in appendix 1A, all 11 Consortia completed the QI Activity Cost Survey. Of those 11 Consortia, one did not have figures on the level of QI outputs, so we could not estimate the measures of cost per unit of QI activity. Of the remaining 10 Consortia, five had the most complete information and were able to allocate costs across the set of QI activities they provided. We rely on the data for those five Consortia in reporting our preferred estimates below. However, we also provide estimates based on all 10 Consortia for completeness.

Given the limitations noted above, the estimated cost per unit of QI activity presented in this chapter should be viewed as an approximate measure of the economic cost of the set of QI supports implemented in the California Consortia examined in this study. In addition, the figures we present are specific to the five California local Consortia that provided the most reliable cost data and therefore are not necessarily generalizable to other Consortia in California or other states. The variation in the cost per unit of QI activity for any given type of QI support (such as coaching and credit-bearing courses) may reflect differences across each Consortia in the nature of the QI support offered, such as the staff who deliver it, the intensity of the support offered, the scale of the local Consortia, and differential costs of labor and other inputs used for the QI activities, among other factors.

Per Unit Cost of QI Supports

Exhibit 8.1 provides the results from our analysis of the cost per unit of QI activity. For each of the five types of QI supports, we report one or more unit cost measures. The five Consortia with the most reliable cost and output data are the source of the estimates in the first four columns. For each unit cost measure, we first list the number of Consortia reporting cost and output data for that measure. Unit cost measures may not be available because either the Consortia did not offer that type of QI support during the recent fiscal year or they did not report the required output measure. Second, for those Consortia with a unit cost measure, we show the minimum, maximum, and average unit cost value. For completeness, in the final two columns, we also report results for the full set of 10 Consortia with unit cost data. For any given unit cost measure, some Consortia may not have a result because they did not offer that QI support during the reporting year or they did not report the required output measure.

Coaching/Mentoring

Of the five types of QI supports, coaching/mentoring is the most prevalent, offered by all 10 Consortia that provided the information required to calculate cost per unit of coaching/mentoring activity. Exhibit 8.1 reports on two unit cost measures for coaching/mentoring: cost per participant and cost per person-hour (measured as total Consortia coaching/mentoring cost divided by the total number of coaching hours provided across all recipients). Among the five Consortia with the most reliable data, the cost per participant ranges from $1,460 to $7,090, with an average value of $3,400. When all 10 Consortia with valid data are included, the average cost per participant is higher by about $1,500, a pattern that holds more generally, because in most cases the cost per unit of activity is higher on average for the other five Consortia with less reliable cost estimates.

Exhibit 8.1. Cost per Unit of QI Activity: QI Consortia Fiscal Year 2013–14 or 2014–15

	Consortia With Most Reliable Cost Data (Maximum *N*=5)				All Responding Consortia (Maximum *N*=10)	
	Number of Consortia Reporting	Cost per Unit ($)			Number of Consortia Reporting	Cost per Unit ($)
Measure		Minimum	Maximum	Average		Average
Coaching/mentoring						
Cost per participant[a]	5	$1,460	$7,090	$3,400	10	$4,960
Cost per person-hour	4	$180	$890	$550	9	$660
Credit-bearing courses						
Cost per participant[a]	3	$1,010	$1,850	$1,350	6	$3,900
Cost per credit-hour	3	$350	$790	$500	6	$640
Noncredit courses						
Cost per course-hour	3	$130	$1,610	$780	7	$2,240
Cost per participant[a]	4	$30	$740	$310	7	$460
Cost per person-hour	3	$40	$370	$160	5	$140
Peer support activities						
Cost per participant[a]	2	–	–	$2,430	2	$2,430
Financial incentives						
Cost per recipient provider	2	–	–	$10,090	6	$6,540

SOURCE: Authors' analysis of the 2015 QI Activity Cost Survey. See appendix 1A for further details on survey methods.

NOTE: Dollar figures are rounded to the nearest $10. "–" = suppressed because this information was available for only two Consortia.

[a] Participant counts are unduplicated.

This per-participant cost measure does not account for variation across Consortia in the number of coaching/mentoring hours each participant receives. Thus, exhibit 8.1 also shows the cost per person-hour of coaching/mentoring, an estimate that is available for four of the five Consortia

with the most reliable data.[29] The cost ranges from just under $200 per person-hour to nearly $900 per person-hour, with an average of $550 per person-hour. Again, the average cost is higher (by about $100 per person-hour) when calculated for all nine Consortia for which a cost per person-hour can be computed. The variation across Consortia in the cost per person-hour for coaching/mentoring may reflect differences in the qualifications of the staff providing the coaching/mentoring (and hence their compensation), as well as differences in other factors that affect cost, such as the caseload for each coach/mentor and the administrative cost for supporting the coaching/mentoring program.

Credit-bearing Courses

Three of the five Consortia with the most reliable cost information reported costs and outputs associated with supporting early learning staff to participate in credit-bearing courses at colleges and universities. For this group of Consortia, the cost per participant receiving this type of QI support was $1,350 on average, ranging from just over $1,000 to almost $1,900 (see exhibit 8.1). Among the 10 Consortia with complete information, six provide support for credit-bearing courses, with an average cost per participant of just under $4,000.

Because Consortia may support varying amounts of course taking, the survey also collected information on the annual course credit hours supported by the Consortia. This allowed us to estimate the cost per credit-hour. For the same three Consortia with the more reliable cost information, there is again a range of costs, from $350 per credit-hour to $790 per credit-hour, with an average of $500 per credit-hour. For the total of six Consortia, with varying degrees of cost data quality, which support credit-bearing courses, the average cost is somewhat higher, at $640 per credit-hour.

Noncredit Courses

Overall, six of the 10 Consortia with complete cost and output information reported costs associated with noncredit courses. Among the five Consortia with the most reliable cost data, the number reporting the required output information to calculate the unit cost measures ranged from three to four Consortia. Exhibit 8.1 shows results based on three unit cost measures: cost per noncredit course-hour (that is, total costs divided by total hours of course offerings), cost per participant (that is, total costs divided by the total number of unique persons attending courses), and cost per person-hour (that is, total costs divided by the total hours of courses taken by all participants). On average, the cost per course-hour is almost $800, the cost per participant is just over $300, and the cost per person-hour is $160. This pattern reflects the efficiency in delivering a single course to multiple participants, as compared with coaching/mentoring, which is typically a one-on-one activity, for each available hour. It is evident, however, that there is considerable variation across Consortia in these different measures of unit cost, both for the five Consortia with the most reliable data, as well as the other Consortia with cost data. This may reflect the

[29] One Consortium did not report the number of annual hours of coaching/mentoring provided by the Consortium, which is required for this unit cost measure.

nature of the noncredit course offerings (for example, who delivers the course and their compensation), as well as the number of attendees per course.

Peer Support Activities

Two Consortia provided information on costs and outputs for peer support activities, which captured such QI activities as provider quality network meetings, provider association support, and student support networks. As both of the reporting Consortia were in our more reliable group, exhibit 8.1 shows the average annual cost per participant of just over $2,400. (The minimum and maximum are not being disclosed, given that only two Consortia reported this information.)

Financial Incentives

All but two of the 11 Consortia reported costs associated with providing various types of financial incentives and supports, but many Consortia did not have estimates of the number of providers, classrooms, or program staff receiving financial supports.[30] Of the five Consortia with the most reliable cost data, four reported costs associated with financial incentives, but only two reported the number of providers receiving such supports. For this reason, we mask the minimum and maximum values, but report the average total cost of financial supports per provider during the fiscal year of just over $10,000 (see exhibit 8.1). This figure is inclusive of both the direct financial incentives transferred to providers or their staff and any direct and indirect costs to the Consortia for supporting this form of QI support. This cost figure could be calculated for four of the five Consortia with less reliable cost information. The resulting overall average across the six Consortia with data is about $6,500. The variation in cost for financial incentives reflects differences across Consortia in the extent to which financial incentives are available and the dollar amount of any given financial incentive available to providers in the Consortia.

Summary

Our goal in this analysis was to collect information on the economic costs associated with the main types of QI strategies employed by local Consortia in California. Such information may be of interest in its own right to understand the resources required to provide various QI supports in the context of a QRIS. In addition, such cost information may provide the basis for undertaking a cost-effectiveness analysis to understand which QI activities produce the largest impacts on program quality improvement or children's developmental gains for every dollar spent. For purposes of this study, it was not possible to conduct this type of comparative cost-effectiveness analysis as estimates of QI activity impacts presented in chapter 7 do not show consistent expected impacts across more than one QI activity. Coaching was the only QI activity that showed consistently positive associations with children's developmental outcomes. Thus, we

[30] Of the nine Consortia that reported costs associated with financial incentives, just two Consortia were able to report the number of providers, the number of classrooms, and the number of staff receiving those supports. For this reason, we focus on one unit cost measure: the cost per provider.

could calculate a cost-effectiveness ratio for coaching but not for other QI activities. Without multiple ratios, a comparative cost-effectiveness analysis cannot be conducted.

To our knowledge, our analysis provides one of the first efforts to estimate the cost of various QI activities on a per-unit basis. We provide such estimates in this chapter for up to 10 of the local Consortia. Given the limitations on the cost information provided by the Consortia, the estimated cost figures should be viewed as approximations of the true economic cost. Furthermore, there is considerable variation across Consortia in the resulting cost estimates. We would expect some variation, given that the Consortia approach the QI activities in different ways and face various costs for labor and other inputs in their local community. But some of the variation may reflect the challenge of calculating such costs in a consistent manner. Future research may seek to understand and document the reasons behind the variations in reported costs by Consortia. This would allow for more meaningful cost estimate comparisons across Consortia and would thus ensure the most reliable cost-effectiveness analyses.

Chapter 9. Summary and Conclusions

In this chapter, we review the findings from the Independent Evaluation of the Race to the Top-Early Learning Challenge Quality Rating and Improvement System (RTT-ELC QRIS) in California, including the Validity and Reliability Study (reported on primarily in the half-term report: http://www.cde.ca.gov/sp/cd/rt/documents/airhalftermreport.pdf) and the Outcomes and Effectiveness Study. The chapter also presents some ideas that the state and Consortia may want to consider as next steps in the refinement of the QRIS.

Summary of Findings

In this section, we summarize the findings presented in the preceding chapters as they align with each of the research questions outlined in the beginning of this report (see exhibit 1.2 in chapter 1). The discussion is organized according to three broad themes: (1) system implementation; (2) QRIS ratings, including the evidence of their validity and perceptions of them in the field; and (3) quality improvement (QI) activities and changes in quality and child outcomes.

System Implementation

1. What is the status of implementation of the RTT-ELC QRIS in 2015, and what are the prospects for sustainability?

California is near the completion of its RTT-ELC QRIS pilot, and, as would be expected of a pilot, the system continues to grow and change. For example, the Consortia have done a considerable amount of work implementing the Hybrid Rating Matrix and fine-tuning and modifying it, with some further revisions still under consideration.

The majority of the Consortia had implemented most of their planned QRIS activities as of summer 2015, including reaching their target goals for the total anticipated number of QRIS participants. As of September 2015, there were more than 2,200 sites enrolled in the QRIS across all participating Consortia (QRIS Online Compendium, 2015), which is almost double the number of sites in the previous year, indicating a substantial increase in the number of sites engaged in QI, and a concomitant step toward the goal of providing high-quality care for California's children.

Consortia continued to provide a range of QI supports to participating sites and staff, and to conduct independent classroom assessments as part of the rating process. And although finding, training, and retaining qualified classroom assessors were major challenges in 2014, these challenges seemed to be somewhat reduced (though not eliminated) in 2015, as Consortia found ways to make classroom observations more affordable and manageable, including making modifications to the Environment Rating Scales (ERS) element of the Hybrid Rating Matrix.

During this QRIS pilot, the Consortia have largely focused on conducting quality assessments and QI activities to help improve program quality. Although ratings have been provided internally to providers, most Consortia had not publicized ratings by summer 2015, and hence we were not able to evaluate the impact of publicized ratings on QI and parent access to quality

programs. Many of the Consortia were still working on a plan for rolling out the ratings, though, according to the state Implementation Team, all had publicized ratings by December 2015.

In terms of validation, it is important to reiterate that California's QRIS is not one uniform system. For example, although all of the Consortia use a common, five-tiered Hybrid Rating Matrix, each Consortium can elect to block Tier 2 and make local modifications to elements within Tier 5. This allowance for local modifications, in addition to the changes at the three-point level and five-point level of the ERS element of the rating matrix in May 2015, shed light on elements that particular Consortia deem most important, but they also complicate attempts to validate the system.

Finally, it may be difficult to expand—and possibly even to sustain—the work done as part of the QRIS pilot without changes in the requirements for participation in the system and/or other financial resources to help rated sites improve. Although the CSPP QRIS Block Grant and First 5 IMPACT grants are expected to make large contributions to ECE in California, and the 17 pilot Consortia in particular (a total of more than $60 million to the 17 pilot Consortia from both IMPACT grants and CSPP QRIS Block Grants), it is unclear whether more private providers and voucher providers will participate in the QRIS system without a requirement that they do so. In addition, now that the pilot counties appear to be publicizing ratings, it will be important to assess the extent to which ratings either inform parents or motivate QI. In a voluntary system, the issue of participation might be particularly tenuous for private providers if they are expected to be assessed without receiving sufficient QI supports.

Although there are issues to be addressed, as noted above, the Consortia, within the limitations of the RTT-ELC grant, have accomplished a great deal in the four years since the grant was awarded and the five years since the California Early Learning Quality Improvement System (CAELQIS) Advisory Committee made its recommendations. Many of the QRISs in other states have taken more than a decade to reach full implementation, and even some of the best-known systems, such as North Carolina's, regard refinement of the system as an ongoing effort. The fact that all 58 counties are now participating in some aspects of the QRIS provides tangible evidence of the growing interest in the system.

2. What incentives or compensation strategies are most effective in encouraging QRIS participation?

As described above, the Consortia have met their goals for their targeted number of participating sites. But, as anticipated, the system includes only a small fraction of the licensed centers in the state (approximately 15 percent), and an even smaller percentage of licensed family child care homes (FCCHs; approximately 2 percent). By comparison, according to the online QRIS Compendium, in states such as Illinois, New Hampshire, and Oklahoma, 100 percent of licensed centers and FCCHs participate in the QRIS. Universal QRIS participation reflects the fact that, in these states, obtaining a license automatically awards a program one star. Also, in Wisconsin, participation is mandatory for programs receiving federal and state subsidies; 82 percent of licensed centers and 75 percent of licensed FCCHs participate in the QRIS. Without the authority to require all licensed programs to participate in the pilot or the resources to encourage their participation within the 16 counties, the Consortia have focused on enrolling publicly funded programs and have limited inclusion of many private centers or FCCHs with full fee-paying

parents. Indeed, several Consortia reported facing challenges recruiting these providers. The publicly funded, and especially publicly contracted programs, in many Consortia have a long history of participation in QI initiatives, and there have frequently been financial and other incentives for their participation. The privately funded programs may have less experience with publicly administered QI initiatives and, short of more incentives, be more wary of participation. But these programs may benefit most from participation in the QRIS.

When asked about their reasons for participating in the QRIS, a majority of directors in the system indicated that their sites chose to participate in the RTT-ELC QRIS to improve their program quality. Similarly, when asked about their reasons for participating in particular QI activities, staff cited an interest in self-improvement first and foremost. Staff also cited QI plans and supervisor recommendations as motivating factors, and FCCH providers cited free classroom materials as a reason for their participation in trainings. In addition, financial incentives were cited by three fourths of teaching staff surveyed as a reason for participating, although only one third of staff reported actually receiving a financial incentive (on average, more than $1,000) to promote their participation in QI activities.

Assessing the role of financial incentives in the QRIS is challenging for several reasons. Most of the RTT-ELC QRIS funds to date have been invested in program quality assessment, which may have been especially costly in the start-up phase of the QRIS, or program quality supports, such as coaching, which have been the backbone of prior QI initiatives that served as the foundation for the QRIS. Hence, there have been few QRIS funds available for financial incentives, and the size of the QRIS-specific financial awards has been relatively small. At the same time, many of the programs participating in the RTT-ELC QRIS in the 16 counties also participate in other QI initiatives (for example, Child Signature Program [CSP-1], Preschool for All [PFA], and Los Angeles Universal Preschool [LAUP]) that predate the QRIS and may offer more robust financial incentives. First 5 PFA in some counties and LAUP, for example, have offered financial incentives large enough to provide salary stipends for staff who receive degrees. Thus, it is difficult to separate the role of QRIS-specific financial incentives from the financial incentives offered by these earlier but, for the most part, still ongoing initiatives. Hence, it is possible that our findings regarding the role of QRIS-specific financial incentives actually understate the role of the whole cluster of financial incentives in motivating participation in QI initiatives.

3. How effective have Consortia been at fostering an improved early childhood system to support early learning and quality improvement in their region? To what extent have the local QRISs been used to align initiatives and projects at the local level?

The RTT-ELC QRIS in 16 counties has been built on existing Quality Improvement System (QIS) and QRIS initiatives, such as the First 5 Power of Preschool, PFA, and CSP initiatives, which likely made it easier for the early implementers to participate. Nonetheless, the QRIS has enhanced communication, collaboration, and alignment among QI initiatives and programs in the participating counties. As one important indication of the degree of collaboration promoted by the RTT-ELC QRIS, LAUP and LA County Office of Child Care's STEP program, which had been separate Consortia within the same county, have agreed to merge. Prior to the RTT-ELC QRIS pilot, these two approaches to a QRIS were very different but now they are merged into

one. Doing so amounts to a reduction in duplication of effort and is, according the state Implementation Team, a credit to the local control model.

In addition, another Consortium has established a coaching collaborative through which it convenes all early childhood education (ECE) coaches from the various initiatives across the county on a monthly basis to share resources, discuss each other's coaching practices, and provide some training and support in order to improve their coaching skills. This collaborative served to improve the coaching services as well as map out where the coaching services were available so that overlap could be reduced. Other Consortia noted that RTT-ELC funds expanded their ability to reach a broader range of programs.

One of the primary goals for a QRIS in California, cited by the CAELQIS Advisory Committee, was to help develop a coherent system from the three sets of program standards operating in the state—Title 22 basic licensing standards, more rigorous State Preschool/Title 5 contract standards, and more comprehensive federal Head Start standards. An additional goal cited by CAELQIS was to motivate and enable more programs to meet nationally recognized standards, such as National Association for the Education of Young Children (NAEYC) accreditation criteria. Toward this end, by developing a framework where Title 22 requirements are roughly equivalent to Tier 1, and Title 5 to Tier 3, Head Start to Tier 4, and NAEYC to Tier 5, the QRIS has provided a logical framework for a program to progress from meeting one set of standards to the next. Administrators also report the RTT-ELC effort has helped promote a common language among ECE professionals within and across programs and agencies. Staff are committed to a common vision and are able to work together toward reaching those goals. Of course, the fact that most of the programs participating in the RTT-ELC to date have been Tier 3 or Tier 4 has limited the extent of the progression from the lowest to the highest tiers. But some QRIS administrators think that the introduction of the CSPP QRIS Block Grant and the First 5 IMPACT grant may contribute to alignment among programs; for example, by encouraging more Tier 1 and Tier 2 programs to participate and attempt to move up to meeting higher standards. But other administrators have concerns about whether these new funding systems, programs, and initiatives are sufficient in scope or volume to impact alignment. For example, the First 5 IMPACT grant requires that Consortia include a certain percentage of FCCHs and infant/toddler programs in the QRIS, but neither of these are eligible for CSPP QRIS Block Grant funds that could be used to sustain themselves. Some QRIS administrators also think that there is more work to be done to align reporting requirements for different funding streams and thus reduce administration and implementation costs.

In summary, the RTT-ELC QRIS has contributed to better communication among the participating providers, but realization of the full potential of a QRIS to improve alignment across programs and provide impetus for programs to progress from one set of standards to the next will depend on the extent to which Tier 1 and Tier 2 programs become part of the system, and the degree to which any of the programs are able to rise to Tier 5.

QRIS Ratings: Evidence of Validity and Perceptions in the Field

4. How effective are the California Common Tiers' structure and components/elements at defining and measuring quality in early learning settings?

To understand how effective the QRIS ratings are at measuring quality in early learning settings, we examined the validity of the QRIS ratings for two purposes: to serve as reliable and meaningful ratings to inform parents about program quality, and to differentiate programs according to the quality of program structures and adult-child interactions. We first examined the measurement properties of the rating, including the internal consistency of the rating and how element scores are related to each other and to the overall rating. In doing so, we find relatively weak associations among the different rating elements. Indeed, some pairs of elements have very low correlations; this is true among centers and FCCHs alike. One implication of this finding is that none of the element scores were redundant, indicating that the elements capture different aspects of program quality and each element provides some unique information about the program. Elements with limited variability tend to be weakly related to the overall QRIS rating and to other element scores. Among centers, the Ratios and Group Sizes element, which has limited variation, is weakly correlated with the overall QRIS rating. This does not imply that ratios and group sizes are not important, but rather that most programs in the study sample met higher bars on this element, perhaps because of regulatory requirements for public programs. Both the Ratios and Group Sizes element and the Developmental and Health Screenings element are weakly related to most other elements. Among FCCHs, the Effective Teacher-Child Interactions (CLASS) element—again, limited in variability—is the element most weakly correlated with the overall QRIS rating and with other elements. The Minimum Qualifications for Lead Teacher/FCCH element, on the other hand, is better correlated with the rating for centers and FCCHs. These results indicate that, in this sample of programs in which overall ratings are fairly homogeneous, some elements contribute more than others to the variability in overall rating, and thus the ability of the rating to differentiate programs based on quality.

Internal consistency of the QRIS ratings is low, particularly among centers, which reflects the weak relationships between some pairs of elements. Low internal consistency does not necessarily suggest that the rating is flawed, but rather that the aspects of quality measured in the different elements are not always closely related to each other. These findings confirm that the California QRIS does not represent an overarching construct of program quality that is unidimensional. With a more diverse group of programs in the system, it is possible that we might see more variability among element scores, which in turn might improve the internal consistency and value of those elements not currently contributing a great deal to overall ratings. Nevertheless, the current evidence suggests that the overall rating on its own may not provide parents with sufficient information about the specific aspects of program quality they may be interested in. Disseminating the element scores in addition to the overall rating level will better serve the purpose of providing parents with meaningful information about program quality.

Second, we assess results of analyses examining the relationships between QRIS ratings or element scores and observed quality. Specifically, we examined the extent to which ratings are associated with a program's average scores on other independent measures of program quality. For this study, we used the Classroom Assessment Scoring System (CLASS) and the Program Quality Assessment (PQA) as the independent measures of quality. Although the CLASS also is

included in the QRIS rating, we included it in this study because other research finds that CLASS scores are predictive of children's outcomes (Howes and others, 2008; Mashburn and others, 2008), and furthermore the CLASS is widely used to validate QRIS ratings thus allowing the study results to be compared with other QRIS research. The PQA is an instrument that measures similar constructs to California's QRIS, but is not included in the rating calculation.

Comparisons of the California QRIS ratings against independent measures of program quality reveal some encouraging relationships. First, ratings are significantly and positively related to CLASS total scores in centers, suggesting that the ratings capture effective teacher-child interactions well. It is important to note, however, that differences in average CLASS scores from one rating level to the next are not all statistically significant; the fact that few of the programs included in the pilot system are at the highest or lowest ends of the rating scale (and thus the range of ratings is truncated) may contribute to this pattern. Second, and more specifically, California QRIS ratings are positively related to CLASS Instructional Support scores and PQA Adult-Child Interaction scores among centers with preschool classrooms. This finding is encouraging because the CLASS Instructional Support domain is the one most predictive of children's outcomes (Howes and others, 2008), and the PQA Adult-Child Interaction subscale captures similar behaviors.

The other CLASS domains (Emotional Support and Classroom Organization) and the overall PQA scores are not significantly related to California QRIS ratings in centers. In addition, for FCCHs, relationships with California QRIS ratings were largely positive for PQA scores and mixed for CLASS scores, although conclusions cannot be drawn from these results because of the limited number of FCCHs included in the analysis.

The results provide preliminary evidence supporting use of the ratings to differentiate programs according to the quality of adult-child interactions, but differences between Tiers 3, 4, and 5 are small and further research is needed to determine whether there is better differentiation comparing programs at lower rating tiers (Tiers 1 and 2) and higher tiers (Tiers 3, 4, and 5).

5. To what extent do the graduated elements and tiers correspond to graduated increases in child outcomes, including (but not limited to) children's learning, healthy development, social-emotional health, and school readiness?

We also examined the validity of the QRIS ratings for the purpose of identifying programs with larger gains in child skills over the course of the program year, although the study was limited in its ability to detect such differences because of the limited range in QRIS ratings among the programs participating early in implementation of the system and the short timeframe for the study. First, it is important to note that among programs participating in the QRIS, children had higher average scores on measures of literacy, math, and executive function at the end of the year compared with the beginning of the year. This was true for each QRIS rating level. But results from analyses examining children's developmental outcomes by site rating tier and element scores do not provide strong evidence of predictive relationships with child outcomes. We find few statistically significant relationships between tier rating and measures of children's developmental outcomes that were included in the study. Compared with children in Tier 3 sites, children in Tier 5 sites had slightly higher executive function, on average, as measured by the Peg Tapping task, but mathematics skills are comparable for children across all tier levels, as are

early literacy skills measured on the letter-word identification measure. Like other findings in this report, however, these results are not necessarily conclusive. The limited associations between California QRIS tier ratings and children's outcomes may be explained, at least in part, by the low internal consistency of the QRIS rating because it is difficult to find clear, linear associations with a measure that exhibits low internal consistency. Limited variability in the program sites included in the early implementation of the QRIS may play a role in both the low internal consistency and the lack of relationships with child outcomes, and results could differ with a broader range of programs. In addition, the ratings could be associated with other child outcomes that were not measured in this study.

We also tested associations between the components of the tier ratings and children's outcomes. We found a mixed pattern of associations between the individual rating elements that comprise the California QRIS tier rating and children's outcomes. Some rating elements do predict children's outcomes, although the direction of findings is inconsistent for these elements and other elements have no association with child outcomes. For example, some structural quality indicators—such as Teacher Qualifications, Ratios and Group Sizes, and Director Qualifications—show a limited positive relationship with children's outcomes. The Child Observation rating element exhibits a negative pattern of association with children's outcomes. The other rating elements do not appear to be consistently related to children's developmental outcomes.

We might expect the elements measuring process quality—especially the Effective Teacher-Child Interactions element—to be most predictive of children's outcomes because they are considered valid measures of program quality that are closest to the child's experience. However, there is little variation in children's outcomes across the CLASS element scores, and the only significant finding is that sites receiving five points, or one or two points, on this element have children who outperform those in sites receiving three points on the Peg Tapping task that measures executive function. However, it is important to note that at one or two points on this element, sites do not receive a CLASS observation. It is possible that these sites would have received higher CLASS element scores if they had been observed. In addition, only 3 centers with a total of 26 children in the sample received a rating of 1 or 2, so estimates for this group are less reliable than for sites that received higher scores on the CLASS element. In fact, the relationships between children's outcomes and the Effective Teacher-Child Interactions element are more consistently positive (though not statistically significant) when the sites with 1 or 2 points on this element are removed. Again, it is important to remember that including a more diverse array of sites in the system may produce different results.

Given the variability in results across the different elements, these findings suggest that the most meaningful information about quality may come from the element scores rather than the overall rating, and the state may wish to provide the element scores to parents in addition to the overall rating. This recommendation is also consistent with the findings of the measurement properties analysis.

Overall, among our limited pool of fully rated sites, we find limited evidence of predictive relationships between the QRIS ratings or element scores and child outcomes. At the early stage of implementation and given the study limitations, it is not surprising that the study does not find strong evidence of predictive relationships with child outcomes. Indeed, previous studies have

found limited utility of associations between QRIS ratings and child outcomes measured during the same year (Elicker and Thornburg 2011; Lahti and others 2014). Future research is needed on the extent to which programs rated at Tier 2 may have different child outcomes than those at higher tiers, since the quality differences may be most meaningful in comparing programs at Tier 2 and at higher tiers. In addition, the state may consider separately reporting those element scores that do relate to child outcomes so parents can consider this factor when using information from the QRIS.

6. To what extent can the Consortia's local QRIS be streamlined and still result in the same program quality level and child outcomes? What common elements of the Hybrid Rating Matrix and Pathways are most important to include?

We also examined how ratings would change under different rating calculation approaches. First, we found that the distribution of rating levels varies substantially by rating approach. The largest changes in the distribution of ratings occur when ratings rely on block designs. Second, we found that element average ratings are more effective than California QRIS ratings at differentiating centers by CLASS and PQA classroom observation scores. Element average ratings are also somewhat better at predicting children's developmental outcomes, although differences are small. And although ratings using blocks are less effective than California QRIS ratings at differentiating centers by CLASS scores, five-level blocks are more effective than California QRIS ratings at differentiating centers, according to the PQA observation scores..

It is important to remember that when interpreting these analyses using alternative rating approaches, they are specific to the sample of centers included in the study. As noted previously, the sample of programs reflects a limited range of ratings and is not representative of the entire population of programs in California. The relationships between alternative rating approaches and observed quality scores may differ for a more diverse group of programs in California.

To identify elements that were most or least important to include in ratings, we examined the element scores and how they relate to the external classroom observation measures as well as children's outcomes.

Perhaps the best validity evidence is for the Effective Teacher-Child Interactions element, where we find positive relationships that also are statistically significant for CLASS and PQA domains, such that sites with higher element scores also have higher scores on these measures of observed quality. In terms of children's outcomes, if we exclude the three sites that received one or two points (which do not require the an actual classroom observation and therefore do not necessarily reflect measured classroom quality), the trends are generally positive, with higher assessment scores for children in sites with higher scores on this element; however, none of these relationships is statistically significant. In addition, the pattern looks different when we include the sites with two points, which tend to have children who score higher than the sites with three points on this element. Given that these sites are not actually evaluated on the quality of classroom interactions, it is perhaps not surprising that the pattern looks a little different there. Nevertheless, the CLASS element appears to be doing a good job of differentiating programs based on quality, at least for those that receive a CLASS observation for this element. And other studies in the research literature find that CLASS scores are associated with children's developmental outcomes (Howes et al., 2008; Mashburn et al., 2008).

Second, we found some evidence for the validity of the Minimum Qualifications for Lead Teacher/FCCH element. When we look at the child outcome data, we see a positive pattern of relationships, such that children in sites receiving more points on this element have higher assessment scores, though only one of these comparisons is statistically significant. We do not find any statistically significant relationships between this element and our quality measures. However, we do see some patterns across our measures. For example, when we consider the CLASS domains, we see a positive (though not statistically significant) relationship between the Minimum Qualifications element and scores on each of the CLASS domains for sites receiving three, four, or five points. Sites that received two points on this element (those with 24 ECE units or an associate teacher permit) do not fit the pattern; they appear to score higher than sites receiving three points, although, again, this difference is not statistically significant. This suggests that adjusting the cut points on this element could potentially improve its validity.

Third, there is the least evidence supporting the validity of the Child Observation element. There are four statistically significant relationships across all of our measures, and they are all negative. There are relatively few sites with one, two, or three points on this element, which limits our ability to find statistically significant relationships, but sites with four points on the Child Observation element appear to score higher on observed measures of quality than sites with five points. The feature that distinguishes the five-point level from the four-point level is a specific aspect of practice related to assessment: the use of the Desired Results Developmental Profile (DRDP) Tech. To receive four points, staff must use the DRDP twice a year to inform curriculum planning; to receive five points, staff also must upload their data into DRDP Tech. It may be that the use of DRDP Tech is not helping teachers to better use or share the assessment data, or it might be that sites that can afford to use this tool (given the technology infrastructure needed) are different in other ways that affect their curriculum planning, assessment, and family involvement practices. The pattern of relationships with child assessment scores is fairly consistent as well, with lower assessment scores at higher point levels on the Child Observation element. Given the definition of the point levels on this element and that the use of the DRDP is so tied to its funding source, this element does not appear to be successfully differentiating programs based on quality and children's outcomes.

7. In context of the findings of the QRIS descriptive study literature review, are there other tiers, resources, measures, tools, or system structures that should be included that support QRIS reliability, validity, and efficiency in program quality and have led to better overall outcomes in other systems or states?

California's Hybrid Rating Matrix covers a broad range of important domains of quality in early care and education settings. Moreover, a research base and precedent exists among other QRISs for the inclusion of the current elements. For example, the Hybrid Rating Matrix includes three of the five most common indicators used across states with QRISs: staff qualifications; program environment; and program administration, management, and leadership. California's QRIS also includes child observations (as do 55 percent of other systems) and teacher-child interaction (as do 48 percent of other QRISs).

However, more than three fourths of QRISs in other states also include curriculum in their rating systems, and a growing number of states require alignment of curricula with state early learning foundations. Although there is not validation evidence to support the inclusion of these elements,

it is noteworthy that it is a priority in so many other states. In addition, the CAELQIS Advisory Committee (2010) recommended aligning curricula with the *California Preschool Learning Foundations,* the *California Preschool Curriculum Framework*, and the *California Infant/Toddler Learning and Development Foundations* as an alternative to recommending a specific list of curricula. However, this recommendation was not included in the Hybrid Rating Matrix, though it is included in the RTT-ELC Continuous Quality Improvement Pathways. Parents participating in focus groups for the study also recognized the importance of a strong curriculum, although they did not know how to evaluate the curriculum themselves, which is the underlying function of a system like the QRIS.

In addition, although 93 percent of state QRISs in 2014 had family partnership as a separate element in their rating system, in California's cross-county system, family partnership is only included as part of the Program ERS element (for example, as a subscale for family involvement on the ERS). The RTT-ELC Continuous Quality Improvement Pathways also includes family involvement, but the elements of this document do not count toward points in the Hybrid Rating Matrix. Nevertheless, parents and providers interviewed for the study highlighted family engagement as an important indicator of program quality.

Other aspects of quality, such as support for dual language learners, cultural competency, and support for children with special needs, also are used in other systems and worthy of attention in the broader QRIS framework and Pathways.

8. How effective are the Consortia in increasing public awareness of the characteristics of early learning program quality that promote better outcomes for children?

Although ratings have been provided internally to providers, most Consortia had not yet disseminated ratings to the public. As of summer 2015, only one Consortium had made ratings available to the public through a searchable online database, both to find programs involved in the QRIS and to review their ratings. Many of the Consortia had concerns about publicizing the ratings of a pilot project, especially given the change in the rating matrix over time, and were still working on a plan for rolling out the ratings during the study timeframe. A few QRIS administrators, on the other hand, were enthusiastic about providing ratings to program sites because it afforded early learning and care providers a tool for improvement and growth. By the end of 2015, according to the state Implementation Team, all of the pilot counties had published ratings.

In developing their plans to publicize the ratings with the public, Consortia focused on how to convey the information clearly to parents and support their understanding of quality. Focus groups with parents indicate that there is still much room for improvement in conveying to parents the meaning and importance of the elements as indictors of quality. Although parents from focus groups generally agreed with the definitions of quality outlined in the Hybrid Rating Matrix, there were some notable disagreements. Although many parents appreciated the value of the Child Observation element, for example, they did not fully grasp the importance of it until after they had enrolled their children in care. Others did not recognize this as a high-priority, quality element. This finding suggests that there is room for additional parent education on the importance of this component of early learning and care program quality. Similarly, the Developmental and Health Screening element was not a priority element for most parents,

although some parents, especially those with children with developmental delays, recognized its importance. Others worried about too much testing of children, which again points to the need for additional parent education to explain that the screening does not amount to testing. Reactions to the Ratios and Group Size element also were somewhat mixed, with some parents highly valuing smaller groups and better ratios, while others were less concerned about (and unfamiliar with the importance of) this indicator. Overall, parents appear to appreciate the above elements of the rating system, but their lack of full understanding of the health and safety requirements also may underscore the importance of outside monitoring of programs to ensure that these requirements are met.

Many parents in focus groups for this study were familiar with their teacher's qualifications and considered them when selecting a program for their child; however, although some felt strongly that teachers should have degrees, most parents did not consider teacher qualifications as the only—or even the most important—indicator of quality. Director qualifications also were generally viewed as important, though parents were largely unaware of their director's qualifications and unsure of how to evaluate them anyway. Similarly, parents' understanding of the CLASS was somewhat mixed—some were aware that their child's teacher was being observed, while others did not; most were unfamiliar with the ERS.

Although some parents expressed different priorities when evaluating the quality of a program, they are very interested in having access to program ratings. In the meantime, parents largely reported relying on recommendations from family and friends as well as online resources to help guide their early care and education choices. For the most part, parents indicated that they want detailed rating information such as element scores rather than a single overall rating. This provides further support for the recommendation, mentioned earlier based on the rating validity evidence, that the state may wish to disseminate both element scores and overall ratings rather than the overall rating alone.

Quality Improvement Activities and Changes in Quality and Outcomes

9. What are early learning staff's experiences with quality improvement activities?

Consortia reported offering a range of QI activities in 2015, and large numbers of staff participating in our survey reported engaging in these activities in the 2014–15 program year. In fact, all staff surveyed reported substantial engagement in QI activities. Moreover, these activities appear to be a regular part of teachers' lives; many teachers in the study sample reported consistent QI activity participation during the school year (although results could differ in a study with a broader range of providers). When asked about the programs providing these QI activities, staff most frequently cited AB212/CARES Plus, Center on the Social and Emotional Foundations for Early Learning (CSEFEL), and Head Start coaches as the program through which they had received coaching support. AB212/CARES Plus also was cited as the top provider of workshops and trainings, followed by the Desired Results Developmental Profile (DRDP) field trainings, CSEFEL, and California Preschool Instructional Network (CPIN). Almost half of staff who reported receiving financial incentives identified AB212 or CARES Plus as the program that provided those incentives.

Of the four activity types they were asked about on the survey, teachers reported the most engagement with coaching and mentoring. They also found coaching and mentoring activities to be the most helpful when compared with activities in the three other categories: noncredit workshops or training, credit-bearing courses, and peer support activities. However, the teachers found all four QI types to be valuable, with more than 80 percent rating each type as helpful. Indeed, coaching was one of the most valued QI activities reported by site directors and QRIS administrators.

When asked to indicate topics in which they would like more support or training, the most mentioned topic area was child behavior management. The next most mentioned was language development/literacy, then additional training on social-emotional development, special needs or inclusion, and math/cognitive development. Most of these topics also were reported as being addressed in staff's current QI experiences, suggesting that the Consortia and the state are focusing their resources where teachers feel they are most needed. The one exception is special needs or inclusion, which was reported as a topic receiving the *least* attention in staff's coaching, training, and peer support activities,

Survey results suggest that staff are actively engaged in QI activities and site leadership sets the expectation for this, with most centers outlining professional development standards to support their QI, including requirements outlined by many sites to receive coaching or mentoring or to participate in noncredit workshops or training. Most center directors also participated in their own QI activities; in fact, their level of participation in coaching was similar to that for teachers.

It is important to remember that these findings represent QRIS survey respondents from the pool of sites with full ratings and may not represent the larger population of early learning sites in the state or even the QRIS. And although we had a high overall response rate (76 percent), it may still be the case that staff members who chose to complete our survey participate in QI activities at different levels or in different ways than those who chose not to complete the survey, thus limiting the generalizability of the results to a wider population.

10. How do the QRIS strategies (for example, technical assistance, quality improvement activities, incentives, compensation, and family and public awareness) improve program quality, improve the professionalization and effectiveness of the early learning workforce, and impact child outcomes? Which strategies are the least and most effective?

As noted above, early learning staff from the sites participating in the QRIS engaged in significant number of QI activities, including credit-bearing courses, thus contributing to their professionalization. In fact, about a fifth of the staff surveyed reported receiving financial incentives for attending college courses, as they work toward a degree. The study examined the relationship between whether teachers participated in any amount of four particular types of QI activities—coaching and mentoring, workshops or training, peer supports, or credit-bearing ECE courses—and both quality outcomes and children's developmental outcomes. (This section addresses participation in *any amount* at all, and Question 11 below examines whether relationships vary by the amount of time in QI activities.)

We found that teacher participation in any amount of coaching was positively associated with children's letter-word identification skills, but not executive function or mathematics skills, however. In addition to positive relationships between coaching in the current year and several study outcomes, the analyses also found some positive relationships between any participation in coaching in the previous year and children's developmental outcomes. Although the study largely found positive relationships between coaching and child outcomes, the outcome related to Story and Print Concepts was an exception. Study analyses found a negative association between participation in any coaching and child scores on this measure. This finding may be an indication that coaching deemphasized these kinds of basic book knowledge skills (for example, identifying the author and title of a book). We also did not find significant associations between participation in coaching and teachers' CLASS scores. As will be discussed in the next section, we observed a more consistently positive relationship between the *amount*, or dosage, of coaching and quality and child outcomes.

For peer supports, there was an observed positive relationship between any participation in this type of activity and all Pre-K CLASS domains; however, that relationship became smaller and nonsignificant in most cases after controlling for prior participation in QI activities and receipt of incentives. The study found no relationship between participation in peer supports and child outcomes.

In terms of participation in ECE coursework, we found no significant relationship between participation in such coursework in the current year and CLASS scores, and a significant negative relationship with literacy outcomes. However, participation in ECE coursework in the prior program year had a positive relationship with these same literacy outcomes. The study design does not allow us to determine if this finding is due to differences between those who did and did not participate in coursework or if there is an alternative explanation for this counterintuitive finding, such as a time lag between the coursework itself and implementation of what was learned to support child learning. For example, those teachers participating in ECE coursework in the current year may actually have found attending ongoing evening and weekend classes somewhat burdensome, thereby in the short run undermining their performance in the early learning programs. In summary, the lack of a relationship between participation in ECE coursework in the current year and child outcomes should not be interpreted as a negative impact of higher education overall, especially because, as noted above, we found some relationship between higher teacher qualifications overall (as measured by the Minimum Qualifications for Lead Teacher/FCCH rating element) and child outcomes. We suggest further research to examine how and when participation in ECE coursework affects child learning.

Finally, we found that attending workshops or trainings in the current program year was not related to child outcomes. Furthermore, participation in training in the prior program year was negatively associated with child outcomes in most models. Although we do not know for certain what might be driving this negative relationship, there may be important, unmeasured differences between teachers who do and not participate in trainings that underlie this unexpected result. There also was no relationship between training and any CLASS outcomes.

These findings, which focus on whether teachers participated in a given type of QI activity at all, are more limited than the results (described below) that consider the level of teachers' participation in QI activities.

11. For which quality improvement activities does increased dosage (time and intensity of participation) impact program quality and child outcomes?

When we consider the intensity of participation in—or "dosage" of—QI activities, whether they be provided within the classroom or elsewhere on- or off-site, we find more consistent relationships with child outcomes in particular. For example, for coaching, we found positive associations between the total hours of coaching teachers received and child outcomes, including letter-word identification, executive function, and early mathematics skills. We also found positive associations between participation in sustained coaching over the year and letter-word identification and executive function. In contrast, as described above, we found that just participating in any amount of coaching was associated with letter-word identification but was not associated with executive function or mathematics skills, suggesting that small amounts of coaching may not be adequate to support teachers in helping children learn in these more challenging areas of child development.

We also found a positive association between the total hours of coaching and the Classroom Organization domain of the Pre-K CLASS instrument, which measures the quality of classroom materials available to children; the time children spend engaged in tasks; and the rules, expectations, and behavior management strategies used by the teacher. The study found some suggestion of a positive relationship between sustained coaching over the year and the emotional support domain of CLASS, which measures teacher sensitivity and the classroom climate, but the relationship was not statistically significant after controlling for prior participation in QI activities and receipt of incentives. In contrast, as noted above, analyses found no significant differences in CLASS scores for teachers who received any coaching versus those who received no coaching, again suggesting that small amounts of coaching may not be enough to support improvements in classroom quality. However, the analyses do not examine cause-and-effect relationships, so more research is needed to determine whether larger amounts of coaching or sustained coaching cause improvements in classroom quality and child outcomes or whether some other explanation, such as teacher motivation, is related to both coaching participation and observed quality and children's outcomes. In addition, further research on the specific coaching models and other critical elements of coaching that best support positive outcomes is warranted.

There was no observed relationship between the intensity of participation in peer supports and CLASS or child outcomes. Although participation in any amount of workshops or training in the current program year was not related to child outcomes, the total hours of workshops or training attended in the current year was negatively associated with executive function.

In all analyses in this study, we cannot identify cause-and-effect relationships because we are unable to account for all meaningful differences between teachers who did and did not participate in different types or amounts of QI—or between children who attend programs in which the teachers receive differing amounts of QI. Although the study analyses control for some teacher and child characteristics, there may be other differences between teachers or children that could explain the observed relationships. In addition, the teachers participating in the study had a relatively high level of education and most taught in highly rated programs with standards-based public funding. Thus, the relationships observed in this sample may not apply to all early childhood teachers in the state or all children in ECE programs in the state. However, these analyses provide initial information about how QI activities are related to child outcomes in

the short term. Clearly, further research—on a broader range of programs, including lower tier programs that have been less exposed to QI activities—is needed to determine the more long-term impact of QI activities.

12. What QRIS strategies and variables best impact measurable site progress through the tiers? What barriers exist in progressing though the tiers?

Because sites participating in the QRIS are rated every two years, it was not possible to examine change over time on tier ratings. Instead, we draw on findings from analyses of QI activities that are associated with changes in quality at the classroom level and reports from early learning staff on the barriers they have faced as they have attempted to move up in their tier rating. As described above, we find some evidence that coaching, especially more intensive participation in coaching, is associated with changes in quality outcomes, which may affect a site's ability to move up a tier level. Moreover, surveys of directors revealed that 8 out of 10 directors of sites at Tier 4 or below were, in fact, working to raise their tier rating. Directors cited a number of barriers to achieving higher ratings, however. For example, the biggest barrier to moving up to the next tier level, reported by 57 percent of directors, was insufficient funding to increase or sustain staff or director compensation to reward increased education levels. About four in 10 sites also reported that completion of required staff education levels (42 percent) and insufficient funding to meet standards or education requirements (39 percent) were major barriers. At the same time, more than half (53 percent) of sites reported that completion of required annual staff professional development training was not a barrier, although results could differ if the study included fewer state-funded sites.

We also asked directors to indicate whether each of seven RTT-ELC QRIS rating elements was especially easy or difficult to attain. Among those elements rated as especially difficult, the highest percentage of sites chose Effective Teacher-Child Interactions: CLASS assessments (32 percent). Only 16 percent to 23 percent of site director respondents described each of the other six elements as especially difficult. It is interesting to note that although only 16 percent of directors highlighted attaining the Minimum Qualifications for Lead Teacher/FCCH as especially difficult, 42 percent reported that completion of the required staff education levels was a major barrier to moving up the tiers. This finding may reflect some lack of clarity on the specific requirements for this element.

13. What is the cost versus benefit for various QRIS strategies relative to child outcomes?

Using information on the economic costs associated with the main types of QI strategies employed by local Consortia in California, we examined the per-unit cost of various QI activities. Among the five Consortia with the most reliable data, the cost per participant averaged $3,400 for coaching and $1,350 for credit-bearing courses. On average, the cost per participant for noncredit workshops or trainings was just over $300 (based on data from four Consortia). This pattern reflects the efficiency in delivering a single workshop to multiple participants as compared with coaching/mentoring, which is typically a one-on-one activity for each available hour. Two Consortia provided reliable data on peer supports; the average annual cost per participant for this type of QI activities was just over $2,400. Thus, coaching, the type of QI activity that was most consistently related to both quality outcomes (that is, CLASS scores) and children's developmental outcomes, also was the most expensive. This is not surprising, given

that, unlike the other types of QI activities, coaching is typically conducted one-on-one and thus is more labor intensive. Peer supports, which were associated with classroom quality outcomes but not child outcomes, was somewhat less expensive, although still much more costly than group-administered workshops, which did not show positive associations with quality or children's developmental outcomes.

Such information may be of interest in its own right to improve understanding of the resources required to provide various QI supports in the context of a QRIS. In addition, such cost information may provide the basis for undertaking a cost-effectiveness analysis to understand which QI activities produce the largest impacts on program quality improvement or children's developmental gains for every dollar spent. For purposes of this study, it was not possible to conduct this type of comparative cost-effectiveness analysis as estimates of QI activity impacts presented in chapter 7 do not show consistent expected impacts across more than one QI activity. Coaching was the only QI activity that showed consistently positive associations with children's developmental outcomes. Thus, we could calculate a cost-effectiveness ratio for coaching, but not for other QI activities. Without multiple ratios, a comparative cost-effectiveness analysis cannot be conducted.

In addition, given the limitations on the cost information provided by the Consortia, the estimated cost figures should be viewed as approximations of the true economic cost. Furthermore, there is considerable variation across Consortia in the resulting cost estimates. We would expect some variation given that the Consortia approach QI activities in different ways and face varied costs for labor and other inputs in their local community. But some of the variation may reflect the challenge of calculating such costs in a consistent manner. Future research might seek to understand and document the reasons behind the variations in reported costs by Consortia. This would allow for more meaningful cost-estimate comparisons across Consortia and would produce the most reliable cost-effectiveness analyses.

Study Conclusions and Limitations

Prior to RTT-ELC, the CAELQIS Advisory Committee proposed evaluating uniform standards of quality in a statewide pilot, but, citing budgetary reasons, the state preferred a more locally driven approach. Using the RTT-ELC grant funds and building on preexisting initiatives in which primarily state- and federally contracted providers participated, the 17 Consortia agreed upon and began conducting ratings under the Hybrid Rating Matrix. Beginning with the sites already engaged in QI efforts, the state also was able to focus its resources on programs that serve at-risk children—a requirement of the grant. However, from the outset, it was clear that there was neither the authority nor the resources to require (or even encourage) the inclusion of a full spectrum of providers in the system. Among those that did participate, many had been participating for years in prior QIS/QRIS efforts, making it difficult to measure the impact of QRIS-specific activities. This context forms the basis for several key limitations that constrain the extent to which firm conclusions can and should be drawn from this study about the validity of the ratings and outcomes associated with participation in the QRIS.

First, the system is new and still in the development and refinement stage. Moreover, in this early stage of implementation from which the study data are drawn, a relatively small pool of sites had full ratings. Of the 1,273 sites in the rating system as of January 2014, only 472 had full

nonprovisional ratings, and, as a result of the initial rollout of the system as described above, these sites differ from provisionally rated sites in terms of funding source and home language use as well as rating—with provisionally rated sites scoring lower on average. Variation in ratings also is truncated. With no sites rated Tier 1 and only a handful rated Tier 2 or 5, it is difficult to assess the extent to which the rating adequately differentiates programs based on quality. With a more diverse pool of sites, the results might look different.

Second, given the smaller pool of fully rated sites eligible for the study and delays with data collection start-up, the sample of sites for analyses using observed quality data was smaller than anticipated. The analyses using observed quality data include those examining QRIS ratings or element scores and observed quality, and those relating QI activities to CLASS scores was smaller than anticipated. Small sample sizes make it difficult to detect small relationships, thus limiting inferences. That is, some analyses might miss potentially significant differences that would be detected with a larger sample size. The sample of FCCHs was especially small, making it impossible to draw conclusions about these programs.

Third, it is important to remember that we are not able to determine causal relationships in any of the analyses in this study because, although we include a number of statistical controls in our models, we are unable to account for all meaningful differences between teachers who did and did not participate in QI activities or who participated in those activities at different levels. We also cannot account for all differences between children who attend programs rated at different levels or in which teachers have different QI experiences. In addition, as noted previously, the teachers participating in the study had relatively high education levels and most taught in highly rated programs with standards-based public funding; thus, the relationships observed in this sample may not apply to all early childhood teachers in the state or to all children in ECE programs in the state. Given these limitations, study findings should not be interpreted to be firm conclusions.

Despite these limitations, however, several key study findings are important to highlight:

1. **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 over 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 disadvantaged children 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, with most sites receiving 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.

2. **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, 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.
3. **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 compared with 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 where children are randomly assigned to programs with different levels of quality. This 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.

4. **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.

Although our study did not show strong relationships between the QRIS ratings and child outcomes, the limited scope of findings from this study is not unexpected and should not be used to conclude that the system is not effective. Most of the programs participating in the QRIS had been participating for years in better funded initiatives, making it difficult to identify the QRIS-specific impacts. Also, there may be something different about the children participating in the lower tier programs—they may represent more fee-paying, and thus likely working parents. Hence, even though they may have attended objectively lower quality programs, their peer group and family background may have provided a buffer for them. So far, there has been neither a requirement nor resources to require or encourage private providers—even those serving publicly funded children—to participate. Until those providers are in the system, we will have only limited information about the system's effectiveness.

Policy Options for Consideration

Although there are limitations to the study results and conclusions should be interpreted in the context of those limitations, our analyses do suggest some directions that may be worth consideration by the state. In this section, we offer some suggestions for modifications to the system or next steps (several of which were also presented in the Half-Term Report) that the state might want to consider in light of the evidence and other contextual factors.

1. *Consider presenting detailed rating information to parents.*

As noted previously, ratings have not, for the most part, been publicly disseminated. From our conversations with parents in the RTT-ELC counties, it does seem that parents are eager for

rating information, and in fact would like more detailed information than just an overall tier rating. Given the multidimensional nature of the Hybrid Rating Matrix and the evidence that the CLASS element best differentiates programs based on observed quality while some of the structural elements better predict children's outcomes, the state might consider presenting some or all element scores or subratings along with summary ratings. This information would enable parents to make finer distinctions between programs that might share the same or similar QRIS rating. The multidimensional nature of the rating and the fact that different rating elements measure different program components indicates that two programs with the same rating may actually have important underlying differences, reflecting varying strengths and weaknesses on different elements. Although the original intent of a hybrid rating system was to provide programs some flexibility in how they could reach certain levels of quality, in practice it makes comparing programs with the same overall ratings problematic—a problem that is easily addressed by providing the element scores in addition to the overall rating. Moreover, parents may value some rating elements more than others; element scores would enable parents to focus their search on programs that rate highest on the elements about which they care most.

In addition to providing element scores or subratings, the state and Consortia should continue to think carefully about how best to present and explain the scores to parents. Focus groups with parents revealed a lack of clarity on some of the elements themselves as well as a lack of understanding of the importance of other quality elements. This suggests a need for additional guidance to help parents understand the value of the quality elements and identify those that are most meaningful to them.

The state may wish to consider conducting a pilot phase around releasing rating data to identify the best communication strategies to ensure that the information is accessible to parents and provides them with the level of detail needed to inform their decisions. A pilot would also be helpful for understanding how parents learn of the ratings, and how they use them to support their decisions about selecting an early learning and care program for their children. For many parents, choices may be so limited they may not have the option to factor the ratings into their decisions. Depending on how parents use the information, it would be important to understand how this impacts program practices and participation in quality improvement activities as well. These questions could be addressed in a future validation phase.

2. *Consider alternative rating strategies to strengthen validity.*

Although some evidence supports the validity of the Hybrid Rating Matrix in its current form, an element average rating approach appears to hold the most promise from among the alternative rating approaches tested. Ratings calculated by taking an average score across elements are more effective than the California QRIS ratings at differentiating centers by CLASS and PQA classroom observation scores. They also are somewhat more effective at predicting children's literacy and mathematics skill development. California's decision makers may wish to consider this approach as a simple alternative to the current rating strategy. Another advantage to this calculation strategy is that it would make it easy to modify the emphasis given to particular elements. For example, by calculating a weighted average, more weight could be given to the Effective Teacher-Child Interactions element since it was the element that best differentiated programs based on quality. However, it is important to remember that these validity results might change when programs with a wider distribution of ratings are included in the analytic sample.

There may be other approaches to consider as well, such as modifying cut points on the element scores (as described above), which might improve the validity of the rating. The data available for the current study did not permit us to evaluate these kinds of refinements, but they may be worth exploring in the future as a way of strengthening the rating to ensure that it is truly differentiating programs based on quality and children's learning outcomes.

3. *Consider a coaching model that offers sustained support over the course of the year.*

Although the study cannot identify cause-and-effect in the relationships between QI supports and observed outcomes for teachers and children, consistent evidence that coaching is related to better classroom quality and more positive developmental outcomes for children suggests that there is some value to providing teachers with ongoing professional development in the form of coaching. Simply having access to a coach once or occasionally, however, may not be enough. More hours of coaching, and, in fact, sustained work with a coach over the course of the year, is what seems to be related to positive outcomes. This is not surprising, as ongoing coaching or mentoring gives teachers the opportunity to get feedback and guidance as they work on implementing new practices in the classroom. Ongoing coaching enables teachers to ask questions, get advice, and refine their practice. This approach to professional development is supported by the research literature as well (Isner and others 2011). Thus, the state and Consortia should continue to consider coaching tied to staff's learning needs as a central QI support available to sites, and perhaps explore ways to offer more sustained coaching opportunities to early learning staff. This should go hand in hand with further research to identify the most effective models of coaching and key elements that are most supportive of positive outcomes among QRIS sites.

4. *Consider exploring the types of peer supports that are available to staff to learn more about effective approaches to this type of QI activity.*

In addition to coaching, we found some limited evidence of the benefits of formal peer support activities for teachers' classroom quality. Peer supports were defined for staff as formal arrangements, such as learning communities, peer support networks, or reciprocal peer coaching arrangements, which were intended to provide opportunities for staff to discuss shared experiences, exchange ideas, and learn from each other. This support could include structured arrangements, such as professional learning communities (PLCs), or less formal, center-based support groups, such as weekly meetings or presentations among site staff or family child care networks. Defined in this way, peer supports represent a broad set of opportunities in which early learning staff may have participated. Given the breadth of this QI support, it would be useful to explore peer support activities in more detail to better understand what staff are actually doing in these collaborative interactions and how they may support their classroom practice. Depending on how these activities are organized, they could mimic the coaching or mentoring relationships described above and provide a less costly option for Consortia, especially as the funding landscape changes. However, it is important to remember that the observed associations between classroom quality and participation in peer support networks were somewhat weak, which may suggest considerable variation in the nature and quality of these interactions. Thus, some attention to what these experiences entail and under what conditions they appear to be more or less effective is warranted.

5. *Consider ways to encourage or require more providers to participate in the system.*

Perhaps the most important issue for the state to consider is how to encourage or require a broader range of programs to be rated and, hence, participate in the QRIS. Although the validation analyses do not directly address this challenge, we frequently note in this report that one of the major limitations of this research has been the relative lack of variation in the sample of programs participating in the study. The majority of programs and providers participating have been at Tier 3 or Tier 4, with no programs rated at Tier 1 and only a few at Tier 2 or Tier 5. Moreover, the sample is heavily skewed toward state- and federally contracted programs that were already held to a set of contract standards intended to focus on quality before the implementation of the RTT-ELC QRIS. The lack of variation in rated quality is not just a problem for researchers attempting to gauge the effectiveness of the system in rating quality; the narrow range of programs participating also limits the potential impact of the QRIS in helping families choose among a large number of rated programs for their young children. It also forgoes an opportunity to assess the quality of the large group of private programs receiving some public funds in the form of vouchers, and makes it difficult for the public or policymakers to determine how best to direct limited resources for quality improvement.

The state might, therefore, want to consider piloting a system in one or more counties that requires all centers and FCCHs receiving any state and federal subsidies to participate in the QRIS. At least nine states require programs receiving subsidies from the federal Child Care and Development Fund to participate in their QRIS, and several states, such as Illinois and Washington, make participation mandatory for school-operated early care and education programs. Another potential benefit of piloting a QRIS that requires participation by all publicly funded providers would be the information that it would give policymakers about the current quality of the programs in which taxpayers are investing, and where and what type of improvements are needed. Finally, such a pilot would provide a more complete picture of the extent to which the rating system captures distinctions between all five tiers in the Hybrid Rating Matrix.

Another option for a pilot would be to automatically award all licensed centers (and possibly FCCHs as well) Tier 1 status with an option to participate in the QRIS, and hence an opportunity to attempt to rise to a higher tier. States such as Illinois, North Carolina, and Oklahoma currently use this approach, and it leads to virtually 100 percent participation of licensed providers in the system. If this approach were adopted on a trial basis in a pilot in one or more counties in California, it would be important to ensure that programs that wanted to participate in the effort to rise to a higher tier had access to program quality assessments so that all sites could be assessed and receive a full—as opposed to provisional—rating. Establishing a process that would ensure such access to newly participating programs would be an important part of a mandatory participation pilot before statewide implementation could be considered.

A variation on the above approach to a next-stage pilot in California would be to focus incentives for participation on the licensed programs newly participating in the QRIS. In effect, financial incentives and QI supports would be directed toward and concentrated on the newly participating programs for the duration of the pilot.

6. *Consider another validation phase once the system is further developed.*

As noted throughout this report, data limitations, due in part to the QRIS's stage of development, constrain the analyses and limit the generalizability of the results. To address this constraint, the state might consider revisiting system validation once refinements currently under discussion are made and once the system is expanded to include a more diverse array of programs. Analyses conducted after additional rounds of ratings have occurred could also examine changes over time in ratings assigned to programs to further explore the effectiveness of QI activities and supports for enhancing the quality of early learning and development programs.

If further analyses are to be conducted, it would be essential for Consortia to collect, maintain, and share with the state additional classroom- and site-level data, and ensure standardization of data collection practices across Consortia such that data are collected and reported consistently. Such data would enable additional analyses and suggest evidence-based refinements; this work would not be possible without these more detailed data. In particular, it would be helpful to have raw element-level data (for example, ratios and ERS scores). In addition to being useful for accountability purposes, retaining these data would permit the examination of element score cut points and the simulation of ratings based on modified cut points in order to refine the element scoring criteria. Such refinements would strengthen the reliability and validity of the ratings, making the QRIS a more meaningful signal of quality for parents and a more effective tool for targeting QI resources.

References

American Institutes for Research and RAND Corporation. 2013. *Local Quality Improvement Efforts and Outcomes Descriptive Study: Final Report.* http://www.cde.ca.gov/sp/cd/ce/documents/localqieffortfinalreport.pdf (accessed January 5, 2016)

Barnard, W., and others. 2006. *Evaluation of Pennsylvania's Keystone STARS Quality Rating System in Child Care Settings.* Pittsburgh, PA: University of Pittsburgh Office of Child Development and the Pennsylvania State University Prevention Research Center.

Barraclough, S. J., and A. B. Smith. "Do Parents Choose and Value Quality Child Care in New Zealand? *International Journal of Early Years Education* 4 (1996): 5–26.

Boller, K., and others. 2010. *The Seeds of Success Modified Field Test: Findings from the Impact and Implementation Studies*. Princeton, NJ: Mathematica Policy Research. http://www.mathematica-mpr.com/publications/PDFs/earlychildhood/seeds_to_success_mft.pdf (accessed December 5, 2015).

Bryant, D. M., and others. 2001. *Validating North Carolina's 5-Star Child Care Licensing System.* Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Center. http://fpg.unc.edu/resources/validating-north-carolinas-5-star-child-care-licensing-system (accessed March 15, 2013).

Burchinal, M. R., and others. "Caregiver Training and Classroom Quality in Child Care Centers." *Applied Developmental Science* 6 (2002): 2–11.

Burchinal, M. R., and others. "Quality of Center C*hild* Care and Infant Cognitive and Language D*evelopment*." *Child Development* 67 (1996): 606–20.

California Department of Education. (2015). *Consortia Implementation Guide for the Tiered Quality Rating and Improvement System (TQRIS).* Sacramento, CA: Author. http://www.cde.ca.gov/sp/cd/rt/rttelcapproach.asp (accessed January 11, 2016).

California Early Learning Quality Improvement System (CAEL QIS) Advisory Committee. 2010. *Dream Big for Our Youngest Children*: *California Early Learning Quality Improvement System Advisory Committee Final Report.* Sacramento, CA: California Department of Education. http://www.cde.ca.gov/sp/cd/re/documents/fnlrpt2010.pdf (accessed April 3, 2015).

Center on the Developing Child. 2007. *A Science-based Framework for Early Childhood Policy: Using Evidence to Improve Outcomes in Learning, Behavior, and Health for Vulnerable Children.* Cambridge, MA: Harvard University. http://developingchild.harvard.edu/resources/reports_and_working_papers/policy_framework/ (accessed April 3, 2015).

Clarke-Stewart, K. A., and others. "Do Regulable Features of Child-care Homes Affect Children's Development?" *Early Childhood Research Quarterly* 17, no. 1 (2002): 52–86.'

Cryer, D., and M. Burchinal. "Parents as Child Care Consumers." *Early Childhood Research Quarterly* 12 (1997): 35–58.

Cryer, D., W. Tietze, and H. Wessels. "Parents' Perceptions of Their Children's Child Care: A Cross-national Comparison." *Early Childhood Research Quarterly* 17 (2002): 259–77.

CSPP QRIS Block Grant. *California State Preschool Program (CSPP) Quality Rating and Improvement System (QRIS) Block Grant.* (2015). http://www.cde.ca.gov/sp/cd/op/csppqrisblockgrant.asp (accessed January 15, 2016).

Curby, T. W., K. J. Grimm, and R. C. Pianta. "Stability and change in early childhood classroom interactions during the first two hours of a day. *Early Childhood Research Quarterly.* 25 (2010): 373-84.

Diamond, A., & C. Taylor. (1996). "Development of an aspect of executive control: Development of the abilities to remember what I said and to 'Do as I say, not as I do.'" *Developmental Psychobiology* 29, 315–34.

Elicker, J., and others. 2011. *Evaluation of Paths to QUALITY, Indiana's Child Care Quality Rating and Improvement System: Final Report*. West Lafayette, IN: Purdue University.

First 5 California. 2015. *First 5 IMPACT Request for Application.*

Goelman, H., and others. "Towards a Predictive Model of Quality in Canadian Child Care Centers." *Early Childhood Research Quarterly* 21, no.3 (2006): 280–95.

Gormley, W. T., and others. "The Effects of Universal Pre-K on Cognitive Development." *Developmental Psychology* 41, no. 6 (2005): 872–84.

Gormley, W. T., and T. Gayer. "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program." *Journal of Human Resources* 40, no. 3 (2005): 533–58.

Helburn, S. W., (ed.) 1995. *Cost, Quality, and Child Outcomes in Child Care Centers: Technical Report*. Denver, CO: University of Colorado at Denver, Department of Economics, Center for Research in Economic and Social Policy.

Helburn, S. W., J. R. Morris, and K. Modigliani. “Family Child Care Finances and Their Effect on Quality and Incentives.” *Early Childhood Research Quarterly* 17 (2002): 512–38.

HighScope Educational Research Foundation. 2003. *Preschool Program Quality Assessment, Second Edition (PQA) Administration Manual.* Ypsilanti, MI: HighScope Press.

HighScope Educational Research Foundation. 2009. *Family Child Care Program Quality Assessment (PQA) Administration Manual*. Ypsilanti, MI: HighScope Press.

Hohmann, M., S. Lockhart, and J. Montie. 2013. *Infant-Toddler Program Quality Assessment (PQA) Form A: Observation Items.* Ypsilanti, MI: HighScope Press.

Howes, C. “Relations Between Early Child Care and Schooling.” *Developmental Psychology* 24 (1988): 53–7.

Howes, C., and others. “Ready to Learn? Children's Pre-academic Achievement in Pre-kindergarten Programs.” *Early Childhood Research Quarterly* 23, no. 1 (2008): 27–50.

Isner, T., K. Tout, and others. 2011. *Coaching in Early Care and Education Programs and Quality Rating and Improvement Systems (QRIS): Identifying Promising Features.* Washington, DC: Child Trends.

Karoly, L. A. 2009. *Preschool Adequacy and Efficiency in California: Issues, Policy Options, and Recommendations*. Santa Monica, CA: RAND Corporation.

Lahti, M., and others. 2013. *Validation of Quality Rating and Improvement Systems (QRIS): Examples from Four States* (OPRE 2013-036). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://www.researchconnections.org/childcare/resources/26590/pdf (Accessed April 3, 2015).

La Paro, K. M., B. K. Hamre, and R. C. Pianta.2012. *Classroom Assessment Scoring System (CLASS) Manual: Toddler*. Baltimore, MD: Brookes.

Malone, L., and others. 2011. *Measuring Quality Across Three Child Care Quality and Improvement Systems: Findings from Secondary Analyses*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://qrisnetwork.org/sites/all/files/resources/gscobb/2011-09-28%2014:15/Report.pdf (accessed March 15, 2013).

Mashburn, A. J. “Quality of Social and Physical Environments in Preschools and Children’s Development of Academic, Language, and Literacy Skills.” *Applied Developmental Science* 12, no. 3 (2008): 113–27.

Mashburn, A., and others. “Measures of Classroom Quality in Pre-kindergarten and Children’s Development of Academic, Language, and Social Skills.” *Child Development* 79, no. 3 (2008): 732–49.

National Association of Child Care Resource and Referral Agencies. 2011. *We Can Do Better, 2011 Update: NACCRRA’s Ranking of State Child Care Center Regulations and Oversight.* Washington, DC: National Association of Child Care Resource and Referral Agencies. http://www.naccrra.org/publications/naccrra-publications/we-can-do-better-2011.php (accessed April 3, 2015).

National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network (ECCRN). “The Relation of Child Care to Cognitive and Language Development.” *Child Development* 74, no. 4 (2000): 960–80.

National Research Council. 2001. *Eager to Learn: Educating Our Preschoolers.* Edited by B. T. Bowman, M. S. Donovan, and M. S. Burns. Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press, 2001.

NICHD ECCRN. “Child-Care Structure → Process → Outcome: Direct and Indirect Effects of Child-care Quality on Young Children's Development.” *Psychological Science* 13, no. 3 (2002): 199–206.

Norris, D. J., and L. Dunn. 2004. *Reaching for the Stars: Family Child Care Home Validation Study Final Report*. Stillwater and Norman, OK: Early Childhood Collaborative of Oklahoma. http://www.okdhs.org/NR/rdonlyres/11E190FA-7FF2-4166-8F6E-19ED34FE4E0F/0/ReachingForTheStarsFamilyChildCareValidationStudyFinalReport_dcc_05092007.pdf (accessed March 15, 2013).

Norris, D. J., L. Dunn, and L. Eckert. 2013. *Reaching for the Stars: Center Validation Study Final Report*. Stillwater and Norman, OK: Early Childhood Collaborative of Oklahoma. http://www.okdhs.org/NR/rdonlyres/4C0EF19D-6FC4-40C6-8926-A3371B7F4130/0/ReachingForTheStarsCenterValidationStudyFinalReport_dcc_05212007.pdf (accessed March 15, 2013).

Peisner-Feinberg, E. S., and others. “The Relation of Preschool Child-care Quality to Children’s Cognitive and Social Developmental Trajectories Through Second Grade.” *Child Development* 72, no. 5 (2001): 1534–53.

Phillips, D., and others. “Within and Beyond the Classroom Door: Assessing Quality in Child Care Centers.” *Early Childhood Research Quarterly* 15. no. 4 (2000), 475–96.

Pianta, R. C., and others. “The Effects of Preschool Education: What We Know—How Public Policy Is or Is Not Aligned with the Evidence Base, and What We Need to Know.” *Psychological Science in the Public Interest* 10, no. 2 (2009): 49–88.

Pianta, R. C., K. M. La Paro, and B. K. Hamre. 2008. *Classroom Assessment Scoring System* (*CLASS*). Baltimore, MD: Brookes.

QRIS National Learning Network. 2015. *QRIS State Contacts & Map.* http://www.qrisnetwork.org/qris-state-contacts-map (accessed December 18, 2015).

QRIS Online Compendium. 2015. *A Catalog and Comparison of Quality Rating and Improvement Systems*. http://www.qriscompendium.org (accessed October 15, 2015).

Sabol, T., and R. Pianta. 2012. *Improving Child Care Quality: A Validation Study of the Virginia Star Quality Initiative*. Charlottesville, VA: University of Virginia Curry School of Education.

Sabol, T., and R. Pianta. "Validating Virginia's Quality Rating and Improvement System Among State-Funded Pre-Kindergarten Programs." *Early Childhood Research Quarterly* 30 (2014): 183–198. http://qrisnetwork.org/sites/all/files/materials/Validating%20Virginia%E2%80%99s%20quality%20rating%20and%20improvement%20system%20among%20state-funded%20pre-kindergarten%20programs%20(Early%20Childhood%20Research%20Quarterly).pdf (accessed April 3, 2015)

Shen, J., W. Tackett, and X. Ma. 2009. *Second Evaluation Report for Palm Beach County Quality Improvement System.* Palm Beach, CA: Children's Services Council of Palm Beach County, 2009.

Shonkoff, J. P., and D. A. Phillips., (eds.) 2000. *From Neurons to Neighborhoods: The Science of Early Child Development.* Washington, DC: National Academy Press, 2000.

Sirinides, P. 2010. *Demonstrating Quality: Pennsylvania Keystone STARS: 2010 Program Report*. Harrisburg, PA: Office of Child Development and Early Learning, 2010. http://www.ocdelresearch.org/Reports/Keystone%20STARS/Keystone%20STARS%202010%20Evaluation%20Report.pdf (accessed March 15, 2013).

Thornburg, K. R., and others. 2009. *The Missouri Quality Rating System School Readiness Study*. Columbia, MO: Center for Family Policy & Research.

Tout, K., and others. 2010. *Evaluation of Parent Aware: Minnesota's Quality Rating System Pilot: Year 3 Evaluation Repor*t. Minneapolis, MN: Child Trends. https://s3.amazonaws.com/Omnera/VerV/s3finder/38/pdf/Parent_Aware_Year_3_Evaluation_Report_Nov_2010.pdf (accessed March 15, 2013).

Tout, K., and others. 2011. *Evaluation of Parent Aware: Minnesota's Quality Rating System Pilot: Final Evaluation Report*. Minneapolis, MN: Child Trends. https://s3.amazonaws.com/Omnera/VerV/s3finder/38/pdf/Parent_Aware_Year_4_Final_Evaluation_Technical_Report_Dec_2011.pdf (accessed March 15, 2013).

U.S. Department of Education. 2011. *Race to the Top Early Childhood (RTT-ELC) Grant Application*. Washington, DC: U.S. Department of Education. http://www2.ed.gov/programs/racetothetop-earlylearningchallenge/applicant.html (accessed April 3, 2015).

Vandell, D. L., and B. Wolfe. 2000. *Child Care Quality: Does It Matter and Does It Need to Be Improved?* Washington, DC: U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation.

Weiland, C., and others. 2013. "Associations Between Classroom Quality and Children's Vocabulary and Executive Function Skills in an Urban Public Prekindergarten Program." *Early Childhood Research Quarterly* 28, no. 2: 199–209. http://www.sciencedirect.com/science/article/pii/S0885200612001172 (accessed April 3, 2015).

Whitebook, M., Sakai, L., and others. 2006. *California Early Care and Education Workforce Study: Licensed Child Care Centers. Statewide 2006*. Berkeley, CA: Center for the Study of Child Care Employment, and San Francisco, CA: California Child Care Resource and Referral Network.

Wolfe, B., and S. Scrivner. 2004. "Child Care Use and Parental Desire to Switch Care Type Among a Low-income Population." *Journal of Family and Economic Issues* 25: 139–62.

Woodcock, R. W., K. S. McGrew, and N. Mather. 2001. *The Woodcock-Johnson Test of Achievement-Third Edition.* Itasca, IL: Riverside Publishing.

Woodcock, R., W., K. S. McGrew, F. A. Schrank, N. and Mather. 2007. *Woodcok-Johnson III Normative Update.* Rolling Meadows, IL: Riverside.

Zaslow, M., and others. 2010. *Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature* (OPRE 2011-5). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Zellman, G. L., and others. 2008. *Assessing the Validity of the Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child-Care Quality*. Santa Monica, CA: RAND Corporation. http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG650.pdf (accessed March 15, 2013)

Zill, N. and G. Resnick, 2000. *FACES Story and Print Concepts – Version 2.* Rockville, MD: Westat, Inc.

Zill, N., and others. 2006. Head Start Performance Measures Center Family and Child Experiences Survey (FACES 2000) technical report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://www.acf.hhs.gov/sites/default/files/opre/tech2k_final2.pdf (accessed April 3, 2015)