Does the California Quality Rating and Improvement System Predict Child Outcomes?

Deborah Stipek

May 2020

PACE
Policy Analysis for California Education

Does the California Quality Rating and Improvement System Predict Child Outcomes?

Deborah Stipek, Stanford University

Acknowledgements

This report, like all PACE publications, has been thoroughly reviewed for factual accuracy and research integrity, and we thank this publication's reviewers. The author assumes full responsibility for the accuracy of the report contents.

PACE

Executive Summary

The use of the Quality Rating and Improvement System (QRIS) to improve early childhood education program quality is based in part on assumptions that the quality of programs can be measured and that quality ratings are associated with meaningful differences in learning outcomes for children. This report reviews all of the state QRIS validation studies that examined associations between individual rating elements and child outcomes as well as other research that exists on four elements of the California QRIS, referred to as Quality Counts California: teacher qualifications, program environment, teacher–child interactions, and child-to-teacher ratio and group size.

The review indicates that the current elements in Quality Counts California have weak and inconsistent associations with child outcomes. The problem is not necessarily with the dimensions measured but with how they are measured and with how points are allocated.

There are a number of methodological reasons that could explain the weak associations found between teacher qualifications and child outcomes. Studies do not always differentiate programs serving different-aged children. Most studies were conducted previous to the increased expectations for preschool teachers' skills and responsibilities. And studies do not examine the quality and content of courses or interactions between preparation and the support early childhood educators experience in their jobs.

The current measures of program environment (ECERS) and teacher–child interactions (CLASS) may be weakly and inconsistently predictive of child outcomes because they do not assess practices that are directly related to the child outcomes assessed. Furthermore, if the current measures are used, greater validity might be achieved by adding points in the QRIS rating with smaller increases in scores at the top of the CLASS scale and at the bottom of the ECERS scale. The evidence suggests similarly that greater validity might be achieved for the child-to-teacher ratio and group size element by awarding more points for variations in the lower end (e.g., under 7.5 for child-to-teacher ratio and under 15 for group size) than for the same differences at the higher end.

The report recommends: (a) developing a program observation measure that is better aligned to desirable child outcomes, including foundational academic skills; (b) creating a measure of teacher qualifications that includes the nature and extent of courses in early childhood education (ECE); (c) differentiating between programs that serve children of different ages; and (d) considering additional ratings related to the quality of work environment and pay. An examination of cutoff scores in rating rubrics and the use of more comprehensive measures of child outcomes in validation studies are also recommended.

Introduction

A primary lever for improving quality in early childhood education (ECE) programs is the Quality Rating and Improvement System (QRIS). QRISs are designed to assess program quality and provide support for improvement. Early childhood programs receive a quality rating based on a set of quality indicators and resources are provided to support quality improvement efforts.

In 2011, California was awarded a federal Race to the Top–Early Childhood Education grant providing $75 million for the development of a statewide QRIS. In contrast to most other states, California opted to implement its QRIS locally, citing the state's diversity, size, and ongoing local quality efforts as rationale. From 2013 to 2016, California engaged in a pilot phase for QRIS implementation in 16 counties. California's QRIS is now being implemented on a voluntary basis in all counties across the state.

California's QRIS is overseen by the California Department of Education and First 5 California but is implemented at the county or regional level through locally operated QRISs. Additionally, there are 10 regional hubs encompassing all 58 counties; this allows for coordination of California QRIS implementation among counties. All counties agree to adopt a common QRIS framework and rating system but are given discretion to make certain local determinations within the common framework.

The county consortium agreed-upon rating matrix consists of seven elements (five for Family Child Care Homes [FCCHs]) organized into three core areas—Core 1: Child Development and School Readiness; Core 2: Teachers and Teaching; and Core 3: Program Environment. The seven elements include child observations; developmental and health screenings; minimum qualifications for lead teachers; effective teacher–child interactions; ratios and group sizes; program environment; and director qualifications. Based on a point system, programs are designated as being in one of five tiers (see Appendix).

Participation in QRIS by early care programs is voluntary; programs serving low-income populations were prioritized during the piloting phase. As of September 2019, California had 10,576 daycare centers and 2,153 infant centers for a total of 12,729 licensed centers in the state. Of these, 4,068—31.9 percent—participated in QRIS. Of the 28,313 FCCHs, 2,661—9.3 percent—participated (California Department of Education, 2020).

The use of the QRIS to improve quality is based in part on assumptions that the quality of early childhood programs can be measured and that the quality represented by different scores is associated with meaningful differences in learning outcomes for children. Evidence for associations between QRIS ratings and children's social and cognitive outcomes is weak. A careful analysis of common elements of QRISs might,

however, provide evidence on which dimensions of program quality are most predictive of child outcomes. This report reviews all of the state QRIS validation studies that examined associations between individual elements and child outcomes as well as other research that exists on four of the elements of the California QRIS. It also discusses measurement issues that may be limiting the value and the strength of associations with child outcomes for particular elements. The first section reviews QRIS validation studies and the second section reviews other research on four of the elements.

QRIS Validation Studies

Early Studies

Between 2008 and 2012, in seven of the studies that were conducted to examine how well QRIS was functioning researchers examined associations between QRIS ratings and child outcomes. Three of these studies did not assess children's development in the fall. As a consequence, any association between QRIS ratings and children's development in the spring could be explained by other variables (e.g., more educated parents choosing higher quality programs). Studies in the other four states, however, did include a fall pretest.

The findings for these four states were mixed (see Karoly, 2014). The study of Colorado (Zellman et al., 2008) found no significant associations between child outcomes and QRIS ratings or elements of QRIS. In Minnesota, QRIS ratings were not significantly associated with child outcomes (Tout et al., 2011) but there were a few scattered associations for specific subscores. Out of 37 analyses examining associations between Early Childhood Environment Rating Scale-Revised (ECERS-R) and Classroom Assessment Scoring System (CLASS) subscores, three were significant in the expected direction and one was significant in the opposite direction.

The study of the Missouri QRIS (Thornburg et al., 2009), in contrast, found that QRIS ratings were significantly associated with children's social skills, measured by the Devereux Early Childhood Assessment (DECA). The effect was explained primarily by two subscales: "initiative" (motivation, persistence) and "self-control" (Thornburg et al., 2009). Additional findings for children living in poverty showed significant associations between QRIS ratings and vocabulary skills; findings for children not living in poverty showed significant associations between QRIS ratings and children's positive relationships with adults.

The study of the Virginia QRIS also showed a few scattered significant associations between QRIS tiers and alphabet knowledge and phonological awareness (Sabol & Pianta, 2015).

Sabol and colleagues (2013) assessed associations between common elements in states' QRISs and child outcomes using data from the National Center for Early Development and Learning (NCEDL) Multistate Study of Pre-Kindergarten and the Study of State-Wide Early Education Programs (SWEEP). They created a composite of the elements to roughly mirror QRIS ratings; these QRIS-like scores predicted prereading skills in only two of the nine states and social skills in only one of the states. Of the specific elements assessed (staff qualifications, including teacher and director level of education and years of experience; staff-to-child ratio and group size; family partnerships; learning environment measured by the ECERS-R; and interactions measured using CLASS), only the CLASS consistently predicted prereading, language, math, and social skills.

Recent Studies

In 2011, California submitted a successful Race to the Top–Early Learning Challenge (RTT–ELC) grant application to the U.S. Department of Education. States that received federal RTT–ELC funds were required to evaluate the validity of their QRIS. Seven states, including California, conducted pre–post child assessments in their evaluation and were thus able to assess associations between QRIS ratings and changes in children's social and cognitive skills (Karoly et al., 2016; Magnuson & Lin, 2016; Maxwell et al., 2016; Quick et al., 2016; Soderberg et al., 2016; Tout et al., 2016; Wellesley Centers for Women & UMass Donahue Institute Applied Research & Program Evaluation, 2017).

Table 1 summarizes the measures used to assess children on various developmental dimensions. Language and literacy were assessed more comprehensively—with at least two different assessments used in every state—than was math. The Woodcock Johnson Applied Problems subtest was used to assess math skills in all but one of the states. Overall, there is considerable overlap in child assessments used by the seven states.

Table 1. Child Assessment Instruments Used in Validity Studies

	Language & Lit	Math	Executive Functions	General Cognition/ School Readiness	Physical Development	Social Emotional	Learning-Related Behavior
California	WJ–Letter Word ID Story & Print Concepts	WJ–Applied Problems	Peg/Pencil Tapping				
Delaware	PPVT WJ–Letter Word ID	WJ–Applied Problems	HTKS			DECA–Protective Factors & Behavior Concerns	
Massachusetts	PPVT WJ–Letter Word ID	WJ–Applied Problems				DECA–Total & Subscales (initiative, self-regulation, attachment, negative behavior)	PBLS Total
Minnesota	TOPEL (print knowledge, phonological awareness) Picture Naming Subtest IGDI	WJ–Applied Problems	Peg/Pencil Tapping	Bracken School Readiness Assessment	BMI	SCBE	PLBS (attention/ persistence)
Rhode Island	WJ–Letter Word ID WJ III–Picture Vocabulary	WJ–Applied Problems	Peg/Pencil Tapping			SCBE	PLBS
Washington	WJ–Letter Word ID PPVT Mullen Expressive Language Early Writing Assessment Pre-LAS	TEAM	HTKS TOQ	Mullen Scales of Early Learning	Mullen Fine & Gross Motor Scales (infants & toddlers)	CBCL	
Wisconsin	WJ–Letter Work ID TOPEL (phonological awareness)	WJ–Applied Problems	HTKS	Bracken School Readiness Assessment		SCBE	PLBS

PPVT: Peabody Picture Vocabulary Test

HTKS: Head, Toes, Knees, Shoulder

DECA: Devereaux Early Childhood Assessment

TOPEL: Test of Preschool Early Literacy

SCBE: Social Competence and Behavior Evaluation (social competence, anger/aggression, anxiety/withdrawal)

PLBS: Preschool Learning and Behavior Scale

IGDI: Individual Growth and Development Indicators

TEAM: Tools for Early Assessment in Math

CBCL: Child Behavior Checklist

BMI: Body Mass Index

TOQ: Task Orientation Questionnaire

Table 2 identifies the states in which significant associations between child outcomes and QRIS ratings and elements were found. In all of these analyses, pretest scores and gender were held constant; in some states, other variables known to be associated with the developmental outcomes assessed (e.g., family income, parent education, and child age) were also held constant. The QRIS elements are defined and assessed in different ways but tend to cluster into the categories shown in the table. Those that did not fit neatly into the categories are explained in the table footnotes. Empty cells occur when the association was not assessed (in most cases because the variable was not included in the study) or because the association was not significant in any state. State validation studies varied in the degree to which they differentiated early childhood settings (e.g., family vs. center) and age groups (e.g., infants and toddlers vs. preschoolers) in their analyses. When findings are reported separately, the differences are noted in the footnotes.

The first row in Table 2 shows the states in which QRIS ratings were significantly associated with the child outcome indicated. Note that the associations were not always in the positive direction and in a few cases were found for one group of children but not another. Associations were not necessarily linear, with each point increase in rating being associated with an increment in the child outcome score. For example, in the case of Massachusetts, children in classrooms with the highest QRIS rating had higher Peabody Picture Vocabulary Test (PPVT) scores than did children in classrooms with the middle QRIS rating, but not higher than those of children in classrooms with the lowest QRIS rating.

In every state, QRIS ratings were significantly associated with at least one assessed child outcome. In the case of California, QRIS ratings were significantly associated only with gains in executive function scores. QRIS ratings for two states—Rhode Island and Minnesota—were significantly associated with three different child outcomes. Looking across states, however, despite some significant associations, as Tout et al. (2017) concluded, "QRIS ratings were not strongly associated with patterns of children's growth across the range of developmental skills assessed in the validation studies" (p. 52).

A few state validation studies conducted additional analyses altering the method for computing the overall rating, for example by using different cutoff scores for allocating points on particular elements. In some cases, these additional analyses yielded a slightly different pattern of significant associations. For simplicity, the table shows the original analyses. The fact that modest changes in how the ratings were computed yielded somewhat different findings, however, suggests the value of investigating the relative validity of different computational strategies.

The other rows in Table 2 summarize the significant associations between child outcomes and scores on individual elements of QRIS.

Table 2. Significant Associations with Child Outcomes Found for States, by QRIS Rating and Elements

	Language & Lit	Math	Executive Functions	General Cognition/ School Readiness	Physical Development	Social Emotional	Learning-Related Behavior
Overall QRIS Rating	~MA | +WA | +MN[1]	+RI[2]	+CA | +DE		+WA[3]	+MA[4] | +MN[1] +RI[5]	+MA[6] | +WI +MN[7] | +RI[5]
Child Health/ Wellness			+CA				
Curriculum	–RI	+RI				+RI	
Teacher Interaction	+WA[8]		–CA | +MN[10]		+WA[9]	+MN[10]	
Administration & Management	–RI[11]	+RI[11]	+DE[12]			+RI[11]	
Environment	+MN[13] | +MN[14] | ~CA[15]		–CA[15]			+MN[14] | +WI[16]	
Ratio and Group Size		+CA | +RI					
Family & Community	–RI | +WA[17]						
Child Assessment	–CA[18] | –RI | +WA[17]		–CA[18]				
Director Qualifications	+CA[19] | +DE	+DE				+RI[20]	
Teacher Qualifications	~DE | –RI	+DE	+CA			+RI[19]	

\+ Association was significantly positive.

~ Significant using one but not another measure or rating system, or effects not consistent in direction (e.g., rating of 1 & 3 lower than 4 & 5; 3 higher than 4).

– Association was significantly negative.

Notes. Significance limited to $p < .05$ or lower. Massachusetts did not provide evidence related to individual QRIS elements; Wisconsin assessed only ERS element.

[1] effect stronger for low income
[2] higher income only
[3] infants & toddlers
[4] DECA Attachment Subscale, positive association for preschool, negative for toddlers
[5] lower income only
[6] only in revised QRIS ratings
[7] attention/persistence
[8] Instructional Support Subscale
[9] engaged support for learning, toddlers
[10] instructional support, anxiety/ withdrawal
[11] administration & management includes "inclusive classroom practices" (written program philosophy on addressing needs of children with disabilities, staff release time)
[12] administration & management element includes support of staff and low turnover
[13] ECERS-R, low income
[14] ECERS-Extension, language scale
[15] ECERS/Family Child Care Environment Rating Scale (FCCERS)/ Infant/Toddler Environment Rating Scale (ITERS)
[16] Interaction Subscale
[17] preschool
[18] Desired Results Developmental Profile (DRDP)
[19] negative for story and print concepts, positive for letter–word identification
[20] teacher and director qualifications combined into one element

Across the states, all the elements were positively associated with at least one child outcome assessed, although not necessarily in a positive direction. More significant associations with QRIS elements were found for language and literacy than were found for math. Possible reasons are that there were more measures used to assess language and literacy (from two to five) than were used to assess math (always one), and if a significant association was found for any of the language and literacy measures, the effect for that state was included in the table. Also, more attention is typically given to language and literacy than to math in early childhood education programs; this dimension may thus be more affected by program quality.

California stands out as having the most negative associations between QRIS elements and child outcomes. The two observation measures—the teacher–child interaction measure (CLASS) and the environment rating scale (ECERS and its derivatives)—were negatively associated with gains in executive functions. The environment rating scale association with children's language and literacy skills (story and print concepts) was uneven, with programs receiving 1, 2, or 4 points on that element being at about the same level and higher than programs receiving 3 or 5 points. The use of child assessments based on observations (mostly the Desired Results Developmental Profile, DRDP) was negatively associated with children's executive functions; as scores on that element increased, children's performance on the executive function measure decreased. The child observation element was also negatively associated with children's performance on the story and print concepts measure; programs having the lowest scores on the element performed the best on the measure.

The findings of state QRIS validation studies do not paint a clear picture of what aspects of early childhood programs are associated with desirable child outcomes. Associations between overall QRIS ratings, as well as specific elements, and child outcomes are inconsistent. There is, however, additional research, not specifically connected to QRIS, on associations between child outcomes and four of the same program quality indices included in California's and many other states' QRIS that might provide a clearer picture of the most important elements to measure. We examine next findings on how well each of the following elements predict child outcomes: teacher qualifications, teacher–child interactions, classroom environment, and child-to-teacher ratio and group size.

Research on Elements

Teacher Qualifications

Many studies of childcare programs have found positive associations between more education and specialized teacher training in ECE and observed classroom quality (Burchinal et al., 2002; Cassidy et al., 2005; Lin & Magnuson, 2018; NICHD ECCRN, 2002; Phillips et al., 2000; Whitebook et al., 1989; for reviews see U.S. Department of Education, 2010; Whitebook, 2003). A meta-analysis conducted by Manning et al. (2019) reported that of the 72 samples contained in 45 studies assessing associations between lead teachers' or caregivers' education (BA vs. not) and ECERS (or its derivatives), 36 showed positive effects of education.

Studies have varied in the levels of teacher or caregiver education studied and which levels were found to make a difference. Cassidy et al. (2005), for example, found that associate's or bachelor's degrees were associated with higher ECERS scores than high school and some college education. In contrast, Burchinal et al. (2002) and Whitebook et al. (1989) found significantly higher scores on the ECERS for caregivers with bachelor's degrees than for caregivers with associate's degrees.

The evidence supporting education, however, is not consistent. Some studies reported no effects of teacher education level, including a BA degree, on observed classroom quality (Early et al., 2006; Lin & Magnuson, 2018; Mashburn et al., 2008; Phillipsen et al., 1997). In an analysis of seven studies (Early et al., 2007), only two studies showed positive associations between teachers' educational attainment and classroom quality, and one study found a negative association.

Studies that examined children's outcomes also yielded mixed results (see Zigler et al., 2011). Some found that higher levels of teachers' education were linked to better school readiness skills. Burchinal et al. (2002), for example, found that a bachelor's degree in ECE was associated with higher language scores in children; Early et al. (2006) found that children whose teachers held a BA had higher math skills than did children whose teachers held an associate's degree. Loeb et al. (2004) reported that education beyond a high school degree was associated with better cognitive skills in children.

In contrast to these studies that found teacher education effects, in the 2013 study of nine states mentioned above, Sabol and colleagues found that staff qualifications, including teacher and director level of education and years of experience, did not predict any child outcomes in any state (see also Mashburn et al., 2008). Falenchuk et al. (2017) reported mixed findings from a meta-analysis of 39 studies. They reported some positive although weak associations between staff education and children's language

outcomes (vocabulary and letter word identification) but no significant association with math (WJ–Applied Problems). Looking across all studies qualitatively, they concluded that "associations between staff education and childhood outcomes [were] non-existent to very borderline positive" (p. 1).

A few studies suggested that the amount of early childhood content in a teacher's education may be more important than their level of education. For example, in the Phillips et al. (2000) study of center childcare programs, teachers' specific training in ECE predicted higher ECERS scores in infant and toddler classrooms, but teachers' education levels did not. Weaver (2002) found that providers with a degree in ECE or college coursework in ECE were rated higher in the quality of care in family childcare homes than were providers without college coursework in ECE. Snider and Fu (1990) reported that more ECE courses, in combination with supervised practice teaching, were strong predictors of teachers' knowledge of developmentally appropriate practices, but level of education was not. Specifically, they found that the courses that differentiated teachers with high or low scores on a measure of their knowledge of developmentally appropriate practice focused on (a) planning, implementing, and evaluating developmentally appropriate content; (b) creating, evaluating, and selecting materials; (c) creating learning environments; (d) learning about curriculum models; and (e) observing and recording behaviors.

One recent study employed a novel strategy for assessing teacher preparation (Lin & Magnuson, 2018)—that is, to use the registry level in the QRIS ratings for teachers, which included 17 steps that combined information on teachers' education, major, credit-based courses, and credentials taken from reviews of transcripts. Analyses using the registry level as a continuous variable showed that additional levels were associated with an increase on the ECERS-R score. Additional analyses indicated that the positive association between the registry level and classroom quality was mainly driven by the difference between teachers with and without credit-based ECE training or credentials.

Summary and analysis. In both the QRIS validity studies and other studies in the research literature, the evidence on the value of credentials is inconsistent and also correlational, which makes causal conclusions risky. It is not, however, fair to conclude that the extent and nature of teacher preparation does not matter. There are many reasons for the inconsistency in results that have to do with the study methodologies and the timing of the research.

First, studies often combined very different ECE contexts and different age groups. It is possible that the preparation required to be an effective childcare provider for infants and toddlers is not the same as the preparation required for preschool teachers and that mixing various ages and types of settings masks these specific effects.

Second, most of the studies were conducted prior to the increased emphasis on academic skill development, especially in preschool. At the time most studies were conducted, even teachers who had BAs in child development may not have been given instruction in teaching academic skills, in which case the program would not be expected to impact children's reading and math skills. Likewise, if academic skills were not emphasized in the childcare or preschool programs in which graduates taught, they would not have the opportunity to apply any training they received in how to teach those subjects. Math, in particular, is given very little attention in preschools (Stipek & Johnson, 2020), which would explain why teacher preparation might not have affected children's math learning.

Any effort to improve teacher preparation in California needs to consider current academic expectations in addition to social and other important domains of development. A study comparing the academic content of kindergarten in 1998 to that in 2006 showed that the amount of time kindergarten teachers spent on literacy learning activities increased by 25 percent (Bassok et al., 2016). To meet the more rigorous kindergarten standards, preschool teachers need to know how to ensure that children have the foundational academic skills they need at kindergarten entry.

Specific courses in subject matter teaching, which are currently rare in California higher education programs preparing preschool teachers, are needed to affect children's academic learning. Research suggests, for example, that a methods course in the teaching of math may be necessary even for teachers to believe that math should be taught to preschool-age children (Platas, 2015). If teachers are not convinced that math is appropriate for preschoolers, they are not likely to teach it, regardless of whether they have a BA or an AA.

Third, the effect of preservice qualifications on child outcomes is most likely influenced by working conditions—especially the amount of support teachers receive on the job. Preservice preparation may not need to be as extensive if it is combined with apprenticeship experiences or strong mentoring on the job (Gardener et al., 2019). A study in California conducted by Vu and colleagues (2008) indicated that having a bachelor's degree predicted quality in programs with relatively few resources and supports, such as community-based childcare, but did not predict quality in programs with more resources as well as ongoing supports and monitoring, such as state-funded preK.

A fourth reason why extant research provides limited guidance on effective teacher preparation is that most studies did not assess the quality or the nature of the courses taken. For example, studies did not differentiate between face-to-face and online courses, which may affect the quality of learning. Or, as is generally true of early childhood educator programs in California, they may have focused more on developing knowledge

about children than on developing the practical skills needed to promote children's learning and development. Only one study was found that specifically examined the effects of clinical practice (Snider & Fu, 1990), which is considered by many experts to be a crucial component of teacher preparation (Darling-Hammond & Bransford, 2005).

The number of courses taken in ECE may be particularly important, as suggested by a few studies reviewed above. One study found that the number of ECE courses teachers took was the most important predictor of classroom quality (Lin & Magnuson, 2018). The number of ECE courses required for a BA that confers an ECE teacher license varies considerably among states. Studies conducted in multiple states could thus mask important differences in the content in BAs, even within ECE degree programs.

More useful for guiding policies on degree requirements would be evidence on the content of courses and other experiences that support the development of effective teachers.

Whitebook and colleagues (2009) provided excellent suggestions for identifying the central elements of productive teacher education. For example, they suggested differentiating between (a) foundational courses that cover theories of learning and development, the philosophy or history of education, and multicultural education; (b) pedagogical courses that focus on methods of teaching or classroom management in general, or as related to particular subject areas (e.g., how best to teach reading); and (c) content or subject matter courses (e.g., English or math courses).

Pointing out that the quality of preparation goes beyond courses, they proposed three categories of training that should also be considered: (a) academic content that balances child development and pedagogy, and that emphasizes working with children from diverse cultural and linguistic backgrounds; (b) opportunities for practice and reflection through field placements; and (c) program structures that support both students and teacher educators as adult learners.

In summary, if the goal is to have quality ratings that are associated with student learning, extant research is too inconsistent and methodologically limited to support California's current allocation of QRIS points for teacher qualifications. It is not clear that the same requirements are desirable for early childhood educators who interact with infants, toddlers, and preschool-age children. Furthermore, most of the research was conducted previous to the increased expectations of preschool teachers' skills and responsibilities, and studies do not examine the quality and content of courses and interactions between preparation and the support early childhood educators experience in their jobs. If teacher preparation is included as an element in Quality Counts California, the rating scale should include more information, including the number of ECE course

credits achieved and whether the training included supervised practice teaching. A rating scale, perhaps modeled on the one used in Wisconsin, might be considered.

Teacher–Child Interactions: CLASS

In addition to being used in most state QRISs, the CLASS has been used in myriad studies in the U.S. and throughout the world. To complete the CLASS, observers watch classroom interactions in cycles of 15 to 20 minutes and then give ratings between 1–7 on 10 items. The items are divided into three subscales: emotional support, classroom organization, and instructional supports.

The authors of the measure reviewed studies showing significant associations between dimensions of the CLASS and various child social and academic outcomes (Teachstone, 2017). For example, they reported studies in which children in classrooms with higher emotional support were rated higher in social competence, were rated lower in behavior problems, and performed better on math and literacy assessments. The classroom organization dimension has also been linked in a few studies to social, literacy, and math skills. And, in a few studies, instructional supports predicted executive functions as well as literacy and math skills. As mentioned above, Sabol et al. (2013) reported that, of the five individual quality indicators included in the nine states they studied, CLASS scores were consistently the strongest predictor of children's learning. The structural quality measures of staff qualifications, staff-to-child ratio, and family partnership provided weaker and less consistent predictions of children's learning. Likewise, a meta-analysis of 19 studies of center-based care programs found a modest but significant effect of the classroom organization dimension on executive functions and of the instructional supports dimension on social skills.

Despite some findings of significant associations between CLASS scores and child outcomes, the evidence is mixed; many studies found no or inconsistent associations (see Farran & Nesbitt, 2020). An Organization for Economic Cooperation and Development (OECD) summary of studies using the CLASS across many different countries (e.g., in Latin America, Europe, and the United States) concluded that "the associations between staff–child interactions [measured by the CLASS] and early child cognitive and socio-emotional skills and competencies are inconsistent in size … few studies show strong and consistent associations between the commonly studied dimensions (e.g., emotional climate, instructional quality and classroom organisation) and any domain of early childhood skills" (OECD, 2018, p. 28). Furthermore, even when a study yielded significant findings, the associations were typically small and sporadic. One study may have found a significant association between a CLASS dimension and a child outcome; another study will not have found that same association, but will have found a significant association for that same CLASS dimension and a different child outcome, and so on.

There is some evidence that variation in CLASS scores above a certain threshold may be more strongly associated with child outcomes than variation among lower scores. Burchinal et al. (2010) examined associations between scores on the CLASS and child outcomes using data from an 11-state preK evaluation. Their findings suggested that the quality of teacher–child interactions was a stronger predictor of social competence and behavior problems in higher than in lower quality classrooms. In a later study, Burchinal et al. (2016) found that the CLASS instructional supports subscale was a stronger predictor of the PPVT and WJ–Letter Word scores in higher quality classrooms (above a 2.75 cutoff) than in lower quality classrooms. Burchinal et al. (2014) found that classroom organization was positively associated with children's behavior only at the higher range. Similarly, Weiland et al. (2013) reported that associations between the CLASS scores and children's executive functions were more consistently found among the higher quality programs. In a study by Smith et al. (2008), CLASS summary scores for classroom organization, emotional support, and instructional supports were positively associated with the measure of cognitive inhibitory control only in the higher range of quality. (CLASS instructional support scores were found to be negatively related to the measure of cognitive inhibitory control in the lower quality range.)

Summary and analysis. Although there is some evidence for associations between CLASS scores and child outcomes, the significant predictions are typically weak and unsystematic. One might expect the classroom organization dimension to be the best predictor of child behavior and the instructional supports dimension to be the best predictor of literacy and math skills. The evidence does not support such correspondences.

There have been many concerns raised about the reliability of CLASS ratings and the particular strategy used to rate classrooms, for example when and for how long the observations are conducted. These are issues that could be addressed without major changes to the measure itself.

There is a significant limitation of the CLASS, however, that would require a substantial change in the content to address. The CLASS items do not map onto the child outcomes commonly assessed, particularly executive functions and foundational math and literacy skills. Evaluations of curricula indicate that improvement of subject matter is achieved through subject-specific teaching. There is some evidence, for example, suggesting that there are very specific strategies that teachers can use to promote executive functions (Diamond et al., 2007; Raver et al., 2011) and that specific teacher behaviors can undermine their development. A recent study by Phillips (2020) found that the best negative predictor of executive functions was what the author referred to as "red flag" negative behaviors—yelling or cursing at a child, ignoring a child in physical or emotional need, physically redirecting or disciplining a child, and using sarcasm or

eye-rolling. Also, many studies have demonstrated that domain-specific measures of the quality of instruction predict subject matter learning. Evaluations of specific math curricula interventions show gains in math skills (e.g., Clements & Sarama, 2008; Schacter et al., 2016). Evaluations of literacy curricula show that the quality of literacy instruction is related to gains in literacy skills (e.g., Bierman et al., 2008; Farver et al., 2009), and studies also show that curricula specifically designed to support language development (e.g., Neuman & Cunningham, 2009; Wasik & Hindman, 2011) and social-emotional skills (e.g., Fantuzzo et al., 2011; Raver et al., 2011) result in gains in those domains respectively.

The CLASS has only three items related to instruction and they are generic. As a consequence, a program could achieve a high score without specific attention to the development of executive functions, literacy, or math skills. If these are important outcomes, then a classroom observation measure that assesses teaching directly related to these outcomes is needed to assess the quality of teaching.

Practical constraints render difficult the widespread use of an observations measure focused on subject matter teaching that is sufficiently nuanced to predict subject matter learning. First, observers would need some specialized knowledge in subject matter teaching. Second, observations would most likely need to be longer than the time allotted for CLASS observations, especially in full-day programs where a particular subject may not be taught within the observation window. Third, the measure itself would need to be more detailed and thus longer. These challenges are substantial but not insurmountable, especially if observations were conducted less often than Quality Counts California currently requires.

The CLASS may be the best of available measures of teacher–child interactions, but the evidence of its associations with children's development on important dimensions is not sufficient to justify the cost. There are many reasons to invest in the development of an alternative observation measure that is better aligned with the desired effects on children's learning and that could be implemented in California in the future.

Classroom Environment (ECERS, ITERS, FCCERS)

The other classroom observation measure included in most QRISs is the ECERS-R. The ECERS-R includes 43 items organized into seven subscales (space and furnishings, personal care routines, language reasoning, learning activities, interaction, program structure, and parent and staff). The ratings are based on a 2- to 3-hour-long observation supplemented by teacher interviews. Studies have suggested that subscores group into two categories: (a) appropriate activities and materials, and (b) learning interactions (Setdoji et al., 2018). The Infant/Toddler Environment Rating Scale-3rd edition (ITERS-3) is designed

for infant–toddler programs and the Family Child Care Environment Rating Scale-3rd edition (FCCERS-3) is designed for family childcare settings.

The more recent ECERS-3rd edition (ECERS-3) gives more attention to the instructional behavior of teachers, but it has not been used in much research. In a recent study, Hestenes et al. (2019) found that ECERS-3 but not ECERS-R scores were significantly associated with the education levels of teachers and teacher-to-child ratios, suggesting that it might be a more sensitive measure of quality.

The evidence on the associations between the ECERS-R and developmental outcomes for young children is mixed. Some studies have reported weak but positive linear associations with preschoolers' receptive and expressive language skills, applied problem-solving skills, and some indices of social-emotional development (e.g., Early et al., 2006; Mashburn et al., 2008). Other studies have failed to find significant associations between the ECERS-R and these dimensions of children's developmental outcomes (e.g., Sabol & Pianta, 2014; Zellman et al., 2008). A review of research that includes the ECERS-R and child outcomes suggests occasional but inconsistent significant associations. Brunsek et al. (2017) reviewed 73 studies and examined 16 studies in a meta-analysis. Meta-analyses revealed a few (3 out of 17) significant, although weak positive relationships between ECERS/ECERS-R total score and language and positive behavior outcomes. The language–reasoning subscale was also weakly related to receptive language.

As was found for the CLASS, there is some evidence for threshold effects, but in the case of ECERS-R, the lower range of the scale appears to be a stronger predictor of child outcomes. Burchinal et al. (2000) combined data across three large-scale childcare studies and, using the quality thresholds defined by the developers of the ECERS, classified classrooms into poor (scores of 1.00–2.99), average (scores of 3.00–4.99), or good (scores above 5.00) quality groups. Children in the poor-quality group showed more behavioral problems and scored lower on verbal, reading, and mathematical problem-solving assessments than did children whose classrooms were classified as of average or good quality. The average and good groups were not different from each other on child outcomes. In another study, variation in the low end of the scale (scores of 1.00 to 3.40) was associated with social-emotional outcomes, but variation among scores above 3.40 was not (Setodji et al., 2018). In a third study, variation in scores in the middle of the scale, from 3.40 to 4.60, was significantly associated with children's receptive and expressive language and applied mathematical problem-solving skills (Le et al., 2015). In all three studies, variation in the upper end of the scale was not associated with child outcomes. Researchers have proposed that the leveling off effect is a consequence of the ECERS-R's limited focus on teaching processes and practices. The ECERS-3 was designed to address this potential limitation. The ECERS-R has also been criticized for its scoring system, which

requires an observer to give a rating at the first place where all indicators are not checked, precluding programs from getting credit for quality indicators at higher levels.

Summary and analysis. Findings from both the QRIS validity evaluations and other studies suggest weak and inconsistent associations between the ECERS and the child outcomes that have been assessed. The ECERS includes a number of subscales (e.g., furnishings, personal care routines) that would not be expected to have direct effects on children's developmental outcomes, although they may have value in their own right. Also, like the CLASS measure, the ECERS does not assess teacher–child interactions that are directly related to the child outcomes typically measured, such as executive functions and language, literacy, and math skills.

Global measures of quality are rarely combined in the same study with more domain-specific measures of quality, so there is no direct evidence on their comparative value. In the one study that included both a global measure (ECERS) and a subject-specific measure (the literacy activities score of the teacher Behavior Rating Scale), the subject-specific measure was the stronger predictor of children's language and literacy skills (Burchinal et al., 2016). This finding provides further support for the value of subject-specific classroom observations.

In brief, the ECERS is not clearly predictive of child outcomes. If it is used, there need to be other reasons, and they should be sufficiently strong to justify the cost.

Child-to-Teacher Ratio and Group Size

Limits on child-to-staff ratios and group size are ubiquitous in state program licensing standards and are considered a key index of quality. Lower ratios and group sizes are expected to improve child outcomes because they allow teachers to know each child better and to devote more time to the needs of individual children.

A recent review of studies assessing associations between child-to-staff ratios and child outcomes for preschool-aged children revealed few significant associations (Perlman et al., 2017). The review included 29 studies with child-to-staff ratios ranging from 5 to 14.5. Only 2 out of 10 studies reported significant associations, with better child-to-staff ratios related to better cognitive child outcomes. Of the 12 studies that assessed math outcomes, one study reported that low child-to-staff ratios were associated with higher math outcomes, and one other study reported that higher child-to-staff ratios predicted higher math outcomes. Most of the studies assessing language outcomes showed nonsignificant associations with child-to-staff ratios but, as was found for math, those that did yield a significant finding were divided between the better outcome being associated

with low ratios (four studies) versus high ratios (two studies). Only 4 of the 22 studies that examined positive social-emotional outcomes showed a positive effect of lower ratios.

A review of studies internationally (which included some of the same studies in the review summarized above) examined the effects of ratio and group size on both quality assessments (e.g., ECERS and CLASS) and child outcomes (OECD, 2018). According to the authors, "lower ratios and, to a lesser degree, smaller group size were found to be consistently supportive of staff–child relationships across different types of settings. However, the evidence for the relationship between smaller ratios and emerging academic skills, such as early literacy and early numeracy, was unclear" (p. 11). Despite the fairly consistent links between ratios and group sizes and staff–child interactions, these structural indicators were not directly linked to child development and learning. The review of the literature indicated a mixed pattern of associations both across and within countries, and the meta-analysis conducted for this report failed to find significant associations between child-to-staff ratios and early literacy and numeracy skills.

Only one study—the California Staff/Child Ratio Study—involved random assignment of children into groups with different ratio levels (Love et al., 1992). Ratios were not associated with teacher reports related to social-emotional dimensions (antisocial, depressed, attention deficit, and immature/dependent behaviors) after controlling for baseline behavior scores.

One analysis suggested that ratio and group size may only matter below a certain level. Bowne et al. (2017) used a comprehensive database of U.S. early childhood education program evaluations published between 1960 and 2007 to evaluate the relationship between group size, child-to-teacher ratio, and program effect sizes for cognitive, achievement, and social-emotional outcomes. All of the programs were for children aged 3 to 5 years. There were no significant linear relationships between child-to-teacher ratio and group size, entered either individually or in combination, and cognitive and achievement outcomes. Instead, both group size and child-to-teacher ratio showed nonlinear relationships with cognitive and achievement effect sizes. Child-to-teacher ratios were significantly associated with cognitive and achievement outcomes only for ratios of 7.5:1 and lower, and for group sizes of 15 and fewer. No discernible relationship was found among larger group sizes and child-to-teacher ratios within the range included. Results were less clear for social-emotional outcomes due to the small sample size. These findings align with prior studies of older children that indicated that child-to-teacher ratios need to be very low to have effects on children's learning (Chingos, 2012; Cho et al., 2012).

Summary and analysis. The California QRIS validation study showed that children who were in programs at QRIS levels 1 or 2 had a modest but significantly lower score on the math assessment (mean score of 3.7) compared with children in programs that

received 3, 4, or 5 points (mean score of 3.8). A smaller child-to-adult ratio was also associated with higher scores on expressive vocabulary in the Rhode Island validation study. Looking across studies, however, the evidence for associations between child-to-staff ratios is surprisingly weak given the strong belief in their importance in the ECE field. One possible reason for the disparity between beliefs and evidence is that the demands on teachers are multiplied with each additional child, and one additional child who has poor self-regulation can challenge even the best teachers. Moreover, children living below the poverty line, some of whom might have experienced physical or emotional trauma in their homes or communities, and children with special needs require a great deal of attention. English language learners also add to the demands on teachers. Constraining the number of children in a classroom most likely reduces the burden on teachers, whether or not it is associated with child outcomes.

Second, the research needs to be interpreted cautiously. The methods used for assessing ratios varied. In some studies, the number of adults and children in the classroom were counted; others relied on staff reports. Aides, volunteers, and teachers were often not differentiated. Three untrained volunteers in a classroom may create a very favorable child-to-adult ratio but may not contribute much to children's learning. The effect of child-to-adult ratio should also depend substantially on the training of the adults, which is not included in extant research. A teacher, aide, or volunteer who has not been trained to manage a group of children may find it hard to be effective even with a small number of children. A highly trained teacher who is an effective manager and offers productive, engaging learning experiences for children may be able to support learning in a much larger group than a teacher with less preparation. International comparisons can be misleading, but in countries such as France, preschool teachers have at least 4 years of training focused on practice and have been found to be very effective with a child-to-staff ratio of 15:1 and a group of as many as 30 children with one aide (OECD, 2018).

Child-to-staff ratios and training requirements both have major effects on cost. California, as with most states in the U.S., has chosen to spend more on keeping ratios low and less on preparing teachers and providing concomitant pay than most European countries. Research designed to understand the trade-offs between these two variables that contribute to cost would be useful. At present, the research supports including only variations at the low end of child-to-staff ratios and class size in a QRIS.

Conclusion

There may be justifications for particular elements in the QRIS that are not related to whether the element is associated with child outcomes. Ensuring health and safety, for example, is important even though it would not be expected to predict children's

social-emotional development, executive functions, or math and literacy skills. And some dimensions, such as child-to-adult ratio and group size, may be related to the health and well-being of staff, whether or not they are associated with child outcomes.

This report focuses on one criterion for selecting program dimensions to include in a QRIS—whether elements are associated with gains on particular developmental dimensions. The present analysis of the QRIS validation studies and other research suggests that the current elements in Quality Counts California have weak and inconsistent associations with child outcomes. The problem, however, is not necessarily with the dimensions measured, but how they are measured and how points are allocated.

Consider measures of the program environment (ECERS) and teacher–child interactions (CLASS). If current measures are used, the evidence suggests that variance in particular ends of the scale are more strongly associated with child outcomes. This research thus has implications for allocating points. Perhaps points should be added with smaller increases in scores at the top of the CLASS scale than at the bottom, and with smaller increases in scores at the bottom of the ECERS scale than at the top. The same analysis applies to ratio and group size. The evidence suggests that greater validity might be achieved by awarding more points for variations in the lower end (e.g., under 7.5 for child-to-teacher ratio and under 15 for group size) than for the same differences at the higher end.

The measures themselves, however, should also be reconsidered, in part because expectations related to child outcomes have changed, especially in the last decade. Historically, kindergarten was created to teach children how to behave in a school setting and preschool (or nursery school, as it was typically called) was designed primarily to give children opportunities to develop social skills by interacting with peers. The emphasis in kindergarten was on school-related behaviors such as listening to the teacher and following directions, and in preschool on child-initiated and cooperative play.

A glance at the skills in the Common Core kindergarten standards as well as the California Preschool Foundations makes it clear that the goals have changed, especially regarding the teaching of foundational academic skills. Many early childhood educators have challenged the increased emphasis on academics, worried that it will undermine children's motivation to learn and crowd out opportunities to play and to develop social skills. The concerns have merit, given evidence that many preschools are using worksheets and other didactic, developmentally inappropriate strategies to teach early literacy and math. Early introduction to academic skills, however, can be provided in a playful, developmentally appropriate way (Stipek, 2017). But doing so requires teachers to have considerable skills, such as understanding the individual needs of children (including dual language learners and children with special needs); creating and implementing

playful activities that lay the foundation for literacy and math skill development and the development of executive functions and social skills; assessing children's progress towards meeting the standards; and using that information to adapt instruction. A program evaluation system that predicts children's progress must therefore assess the degree to which preschool teachers are prepared to implement effective teaching strategies specific to the desired domains of development and whether they actually implement them in the programs in which they teach. Current licensing standards in California do not require people preparing to become preschool teachers to learn how to teach literacy or math, and neither the ECERS nor the CLASS assess literacy or math teaching.

There are also limitations related to the measures of child outcomes typically used in validation studies. For practical reasons, the measures are brief and fairly superficial. They do not come close to assessing all of the outcomes included in the California Preschool Foundations and brief, superficial measures are not as sensitive to the quality of instruction as are more comprehensive measures. If we want to detect the effects of the quality of programs, we will have to use better measures of children's learning.

Quality Counts California also needs to differentiate between the age of children as well as the learning context (e.g., home vs. center). It is not clear that the same set of elements should be used for caregivers of infants and toddlers as are used for preschool teachers or caregivers of children aged 3 to 5 years. There is overlap, to be sure. For example, there is clear evidence that nurturing, sensitive caregiving supports children's development regardless of their age. But while a language-rich environment (including math language) is important for children at all ages—and even toddlers can begin to learn number concepts (e.g., counting a few objects) and prereading skills—children in the 3–5 year range are ready for more structured (albeit playful) instruction related to literacy and math. These differences in expectations should be reflected in a QRIS.

California should consider including elements that are not currently part of Quality Counts. Some states include elements related to features of the work environment (Whitebook et al., 2018). There is some evidence that pay and benefits are associated with the quality of care (Phillips et al., 2000; see Phillips et al., 2016). Low salaries, for example, are associated with higher rates of teacher stress and depression, which are associated with withdrawal as well as with insensitive and intrusive behaviors towards children (Hamre & Pianta, 2004; Sandilos et al., 2015; Whitaker et al., 2015). Low pay is also associated with teacher turnover, which undermines continuity for children and opportunities for them to form secure attachments with caregivers (Whitebook et al., 2014). Teacher turnover itself could be considered as a measure of program quality, in addition to pay and benefits.

Given the evidence on the positive effects of curricula focused on specific domains of learning, whether an evidence-based curriculum is used for literacy and math in

preschool should be considered. The degree to which teachers are provided training and ongoing support in using the curriculum is, however, just as important as its adoption.

Opportunities to participate in professional development and ongoing support for improvement, such as through coaching, are also associated with quality staff–child interactions and child outcomes (OECD, 2018). Attention, however, needs to be given to the nature of support for quality improvement. One meta-analysis found that interventions that included coaching were up to three times more effective than interventions that included in-service training but no coaching (OECD, 2018). But even within coaching there is wide variation in quality, making it important to implement some kind of quality control system.

Policy Implications

This review focused only on CA-QRIS's ability to predict child outcomes. The policy implications accordingly only apply if that is a central goal of the QRIS. The weak and inconsistent findings of the research reviewed here suggest that substantial work needs to be done before QRIS ratings can be expected to predict children's learning and development consistently. The following strategies are suggested to move Quality Counts California towards a valid assessment of program quality.

1. An investment needs to be made in developing a classroom observation measure that is better aligned with desired child outcomes.
2. Research designed to validate the QRIS needs to include more comprehensive measures of children's literacy and math skills than are currently used.
3. A more nuanced measure of teacher qualifications, including the nature and extent of courses in ECE and supervised practice teaching should be created.
4. California should investigate alternative variables to include in its QRIS that are associated with quality, such as staff pay, staff reports of working conditions, the implementation of research-based curricula, and opportunities for high-quality professional development.
5. Adjustments should be made to ensure the appropriateness of quality ratings for children of different ages and in different settings.
6. Studies should examine different rating rubrics and different cutoff scores for awarding points to determine which strategies are the most predictive of child outcomes. Cutoffs for awarding points should be informed by extant research on the measures used.

References

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open, 2*(1). https://doi.org/10.1177/2332858415616358

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., Blair, C., Nelson, K. E., & Gill, S. (2008). Promoting academic and social-emotional school readiness: The head start REDI program. *Child Development, 79*(6), 1802–1817. https://doi.org/10.1111/j.1467-8624.2008.01227.x

Bowne, J. B., Magnuson, K. A., Schindler, H. S., Duncan, G. J., & Yoshikawa, H. (2017). A meta-analysis of class sizes and ratios in early childhood education programs: Are thresholds of quality associated with greater impacts on cognitive, achievement, and socioemotional outcomes? *Educational Evaluation and Policy Analysis, 39*(3), 407–428. http://dx.doi.org/10.3102/0162373716689489.

Brunsek, A., Perlman, M., Falenchuk, O., McMullen, E., Fletcher, B., & Shah, P. S. (2017). The relationship between the Early Childhood Environment Rating Scale and its revised form and child outcomes: A systematic review and meta-analysis. *PLOS ONE, 12*(6), Article e0178512. https://doi.org/10.1371/journal.pone.0178512

Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science, 6*(1), 2–11. https://doi.org/10.1207/S1532480XADS0601_01

Burchinal, M. R., Peisner-Feinberg, E., Bryant, D. M., & Clifford, R. (2000). Children's social and cognitive development and child-care quality: Testing for differential associations related to poverty, gender, or ethnicity. *Applied Developmental Science, 4*(3), 149–165. https://doi.org/10.1207/S1532480XADS0403_4

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*(2), 166–176. https://doi.org/10.1016/j.ecresq.2009.10.004

Burchinal, M., Vernon-Feagans, L., Vitiello, V., Greenberg, M., & The Family Life Project Key Investigators. (2014). Thresholds in the association between child care quality and child outcomes in rural preschool children. *Early Childhood Research Quarterly, 29*(1), 41–51. https://doi.org/10.1016/j.ecresq.2013.09.004

Burchinal, M., Zaslow, M., & Tarullo, L. (Eds.). (2016). Quality thresholds, features, and dosage in early care and education: Secondary data analyses of child outcomes. *Monographs of the Society for Research in Child Development, 81*(2).

California Department of Education. (2020, January). *QRIS common data file.* https://www.cde.ca.gov/sp/cd/rt/documents/caqriscommondatafile1617.pdf

Cassidy, D. J., Hestenes, L. L., Hegde, A., Hestenes, S., & Mims, S. (2005). Measurement of quality in preschool child care classrooms: An exploratory and confirmatory factor analysis of the early childhood environment rating scale-revised. *Early Childhood Research Quarterly, 20*(3), 345–360. https://doi.org/10.1016/j.ecresq.2005.07.005

Chingos, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate. *Economics of Education Review, 31*(5), 543–562. https://doi.org/10.1016/j.econedurev.2012.03.002

Cho, H., Glewwe, P., & Whitler, M. (2012). Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. *Economics of Education Review, 31*(3), 77–95. https://doi.org/10.1016/j.econedurev.2012.01.004

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443–494. https://doi.org/10.3102/0002831207312908

Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do.* Jossey-Bass.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007, November 30). Preschool program improves cognitive control. *Science, 318*(5855), 1387–1388. https://doi.org/10.1126/science.1151148

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., Howes, C., & Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly, 21*(2), 174–195. https://doi.org/10.1016/j.ecresq.2006.04.004

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., Cai, K., Clifford, R. M., Ebanks, C., Griffin, J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H.-J., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*(2), 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

Falenchuk, O., Perlman, M., McMullen, E., Fletcher, B., & Shah, P. S. (2017). Education of staff in preschool aged classrooms in child care centers and child outcomes: A meta-analysis and systematic review. *PLOS ONE, 12*(8), Article e0183673. https://doi.org/10.1371/journal.pone.0183673

Fantuzzo, J. W., Gadsden, V. L., & McDermott, P. A. (2011). An integrated curriculum to improve mathematics, language, and literacy for Head Start children. *American Educational Research Journal, 48*(3), 763–793. https://doi.org/10.3102/0002831210385446

Farran, D. C., & Nesbitt, K. T. (2020). New information on evaluating the quality of early childhood education programs. In O. N. Saracho (Ed.), *Handbook of research on the education of young children* (4th ed., pp. 333–347). Routledge.

Farver, J. A. M., Lonigan, C. J., & Eppe, S. (2009). Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development, 80*(3), 703–719. https://doi.org/10.1111/j.1467-8624.2009.01292.x

Gardener, M., Melnick, H., Meloy, B., & Barajas, J. (2019, December). *Promising models for preparing a diverse, high-quality early childhood workforce* [Report]. Learning Policy Institute. https://learningpolicyinstitute.org/product/preparing-diverse-high-quality-early-childhood-workforce-report

Hamre, B. K., & Pianta, R. C. (2004). Self-reported depression in nonfamilial caregivers: Prevalence and associations with caregiver behavior in child-care settings. *Early Childhood Research Quarterly, 19*(2), 297–318. https://doi.org/10.1016/j.ecresq.2004.04.006

Hestenes, L. L., Rucker, L., Wang, Y. C., Mims, S. U., Hestenes, S. E., & Cassidy, D. J. (2019). A comparison of the ECERS-R and ECERS-3: Different aspects of quality? *Early Education and Development, 30*(4), 496–510. https://doi.org/10.1080/10409289.2018.1559681

Karoly, L. A. (2014, May). Validation studies for early learning and care quality rating and improvement systems: A review of the literature [Working paper]. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/working_papers/WR1000/WR1051/RAND_WR1051.pdf

Karoly, L. A., Schwartz, H. L., Setodji, C. M., & Haas, A. C. (2016). *Evaluation of Delaware stars for early success* [Final report]. RAND Corporation. https://www.rand.org/pubs/research_reports/RR1426.html

Le, V.-N., Schaack, D. D., & Setodji, C. M. (2015). Identifying baseline and ceiling thresholds within the Qualistar early learning quality rating and improvement system. *Early Childhood Research Quarterly, 30*(B), 215–226. https://doi.org/10.1016/j.ecresq.2014.03.003

Lin, Y.-C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly, 42*, 215–227. https://doi.org/10.1016/j.ecresq.2017.10.003

Loeb, S., Fuller, B., Kagan, S. L., & Carrol, B. (2004). Child care in poor communities: Early learning effects of type, quality, and stability. *Child Development, 75*(1), 47–65.

Love, J., Ryer, P., & Faddis, B. (1992). *Caring environments: Program quality in California's publicly funded child development programs: Report on the legislatively mandated 1990±91 staff/child ratio study.* RMC Research Corporation.

Magnuson, K., & Lin, Y.-C. (2016, March). *Validation of the QRIS YoungStar rating scale: Report 2. Wisconsin early child care study findings on the validity of YoungStar rating for children's school readiness.* University of Wisconsin School of Social Work and Institute for Research on Poverty. https://dcf.wisconsin.gov/files/youngstar/pdf/validationreport2.pdf

Manning, M., Wong, G. T. W., Fleming, C. M., & Garvis, S. (2019). Is teacher qualification associated with the quality of the early childhood education and care environment? A meta-analytic review. *Review of Educational Research, 89*(3), 370–415. https://doi.org/10.3102/0034654319837540

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

Maxwell, K. L., Blasberg, A., Early, D. M., Li, W., & Orfali, N. (2016, December). *Evaluation of Rhode Island's BrightStars child care center and preschool quality framework.* Chapel Hill, NC: Child Trends. http://www.brightstars.org/uploads/RI_BrightStars_Evaluation_Full_Report_2016.pdf

Neuman, S., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532–566.

NICHD Early Child Care Research Network. (2002). Child-care structure→ process→ outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science, 13*(3), 199–206. https://doi.org/10.1111/1467-9280.00438

Organization for Economic Cooperation and Development. (2018). *Engaging young children: Lessons from research about quality in early childhood education and care* [Report]. OECD Publishing. https://doi.org/10.1787/9789264085145-en

Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., & Shah, P. S. (2017). Child–Staff ratios in early childhood education and care settings and child outcomes: A systematic review and meta-analysis. *PLOS ONE, 12*(1), Article e0170256. https://doi.org/10.1371/journal.pone.0170256

Phillips, D. (2020). *First do no harm: How classroom processes support or undermine children's self-regulation across the pre-K year*. Georgetown University, Washington DC.

Phillips, D., Austin, L. J. E., & Whitebook, M. (2016). The early care and education workforce. *The Future of Children, 26*(2), 139–158.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly, 15*(4), 475–496. https://doi.org/10.1016/S0885-2006(01)00077-1

Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly, 12*(3), 281–303. https://doi.org/10.1016/S0885-2006(97)90004-1

Platas, L. M. (2015). The mathematical development beliefs survey: Validity and reliability of a measure of preschool teachers' beliefs about the learning and teaching of early mathematics. *Journal of Early Childhood Research, 13*(3), 295–310. https://doi.org/10.1177/1476718X14523746

Quick, H. E., Hawkinson, L. E., Holod, A., Anthony, J., Muenchow, S., Parrish, D., Martin, A., Weinberg, E., Lee, D. H., Cannon, J. S., Karoly, L. A., Zellman, G. L., Faxon-Mills, S., Muchow, A. N., Berglund, T., & Haggard, M. S. (2016, August). *Independent evaluation of California's Race to the Top-Early learning challenge quality rating and improvement system* [Cumulative technical report]. RAND Corporation. https://www.cde.ca.gov/sp/cd/rt/documents/rttelcqrisevalreport.pdf

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-Regulation as a mediating mechanism. *Child Development, 82*(1), 362–378. https://doi.org/10.1111/j.1467-8624.2010.01561.x

Sabol, T., & Pianta, R. (2015). Validating Virginia's quality rating improvement system among state-funded pre-kindergarten programs. *Early Childhood Research Quarterly, 30*, 183–198

Sabol, T. J., & Pianta, R. C. (2014). Do standard measures of preschool quality used in statewide policy predict school readiness? *Education Finance and Policy, 9*(2), 116–164. https://doi.org/10.1162/EDFP_a_00127

Sabol, T. J., Soliday Hong, S. L., Pianta, R. C., & Burchinal, M. R. (2013, August 23). Can rating pre-k programs predict children's learning? *Science, 341*(6148), 845–846. https://doi.org/10.1126/science.1233517

Sandilos, L. E., Cycyk, L. M., Scheffner Hammer, C., Sawyer, B. E., López, L., & Blair, C. (2015). Depression, control, and climate: An examination of factors impacting teaching quality in preschool classrooms. *Early Education and Development, 26*(8), 1111–1127. https://doi.org/10.1080/10409289.2015.1027624

Schacter, J., Shih, J., Allen, C. M., DeVaul, L., Adkins, A. B., Ito, T., & Jo, B. (2016). Math shelf: A randomized trial of a prekindergarten tablet number sense curriculum. *Early Education and Development, 27*(1), 74–88. https://doi.org/10.1080/10409289.2015.1057462

Setodji, C. M., Schaack, D., & Le, V.-N. (2018). Using the early childhood environment rating scale-revised in high stakes contexts: Does evidence warrant the practice? *Early Childhood Research Quarterly, 42*, 158–169. https://doi.org/10.1016/j.ecresq.2017.10.001

Smith, M. W., Brady, J. P., & Anastasopoulos, L. (2008). *User's guide to the early language and literacy classroom observation pre-k tool.* Brookes Publishing.

Snider, M. H., & Fu, V. R. (1990). The effects of specialized education and job experience on early childhood teachers' knowledge of developmentally appropriate practice. *Early Childhood Research Quarterly, 5*(1), 69–78. https://doi.org/10.1016/0885-2006(90)90007-N

Soderberg, J., Joseph, G. E., Stull, S., & Hassairi, N. (2016, May). *Early achievers standards validation study* [Final report]. Childcare Quality and Early Learning Center for Research & Professional Development. https://www.dcyf.wa.gov/sites/default/files/pdf/EarlyAchievers_Validationstudy.pdf

Stipek, D. (2017). Playful math instruction in the context of standards and accountability. *Young Children, 72*(3), 8–13.

Stipek, D., & Johnson, N. (2020). Developmentally appropriate practice in early childhood education redefined: The case of math. In B. Graue (Ed.), *Advancing knowledge and building capacity for early childhood research: Creating synergies among segregated scholarly communities.* American Education Research Association.

Teachstone. (2017, Spring). *Effective teacher–child interactions and child outcomes: A summary of research on the Classroom Assessment Scoring System (CLASS) pre-k–3rd grade.* https://curry.virginia.edu/sites/default/files/CASTL-Projects/CLASS-PD/CLASS Outcomes Summary.pdf

Thornburg, K. R., Mayfield, W. A., Hawks, J. S., & Fuger, K. L. (2009, October). T*he Missouri quality rating system school readiness study.* Center for Family Policy & Research. https://www.marc.org/Community/Early-Learning/pdf/QRSfindings.aspx

Tout, K., Cleveland, J., Li, W., Starr, R., Soli, M., & Bultinck, E. (2016). *The parent aware evaluation: Initial validation report.* Child Trends. https://www.researchconnections.org/childcare/resources/31492

Tout, K., Magnuson, K., Lipscomb, S., Karoly, L., Starr, R., Quick, H., Early, D., Epstein, D., Joseph, G., Maxwell, K., Roberts, J., Swanson, C., & Wenner, J. (2017, December). *Validation of the quality ratings used in quality rating and improvement systems (QRIS): A synthesis of state studies* [OPRE Report #2017-92]. OPRE. https://www.acf.hhs.gov/sites/default/files/opre/cceepra_qris_validation_report_b508.pdf

Tout, K., Starr, R., Isner, T., Cleveland, J., Albertson-Junkans, L., Soli, M., & Quinn, K. (2011, December). *Evaluation of parent aware: Minnesota's quality rating and improvement system pilot* [Final evaluation report]. Child Trends. https://s3.amazonaws.com/Omnera/VerV/s3finder/38/pdf/Parent_Aware_Year_4_Final_Evaluation_Technical_Report_Dec_2011.pdf

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2010). *Toward the identification of features of effective professional development for early childhood educators* [Literature review]. Washington, DC.

Vu, J. A., Jeon, H.-J., & Howes, C. (2008). Formal education, credential, or both: Early childhood program classroom practices. *Early Education and Development, 19*(3), 479–504. https://doi.org/10.1080/10409280802065379

Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology, 103*(2), 455–469. https://doi.org/10.1037/a0023067

Weaver, R. H. (2002). Predictors of quality and commitment in family child care: Provider education, personal resources, and support. *Early Education and Development, 13*(3), 265–282. https://doi.org/10.1207/s15566935eed1303_2

Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childhood Research Quarterly, 28*(2), 199–209. https://doi.org/10.1016/j.ecresq.2012.12.002

Wellesley Centers for Women & UMass Donahue Institute Applied Research & Program Evaluation. (2017). *Massachusetts quality rating and improvement system (QRIS) validation study* [Final report]. Massachusetts Department of Early Education and Care. https://www.mass.gov/files/2017-08/Revised Validation Study ReportfinalFORMATTED.pdf

Whitaker, R. C., Dearth-Wesley, T., & Gooze, R. A. (2015). Workplace stress and the quality of teacher–children relationships in Head Start. *Early Childhood Research Quarterly, 30*(A), 57–69. https://doi.org/10.1016/j.ecresq.2014.08.008

Whitebook, M. (2003). *Early education quality: Higher teacher qualifications for better living environments. A review of the literature.* Institute of Industrial Relations, University of California, Berkeley.

Whitebook, M., Gomby, D., Bellm, D., Sakai, L., & Kipnis, F. (2009, May). *Preparing teachers of young children: The current state of knowledge, and a blueprint for the future* [Policy report]. Center for the Study of Child Care Employment, University of California, Berkeley. https://cscce.berkeley.edu/preparing-teachers-of-young-children-the-current-state-of-knowledge-and-a-blueprint-for-the-future/

Whitebook, M., Howes, C., & Phillips, D. (1989). *Who cares? Child care teachers and the quality of care in America. Final report of the national child care staffing study.* Child Care Employee Project.

Whitebook, M., McLean, C., Austin, L. J. E., & Edwards, B. (2018). *Early childhood workforce index 2018.* Center for the Study of Child Care Employment, University of California, Berkeley. https://cscce.berkeley.edu/early-childhood-workforce-2018-index/

Whitebook, M., Phillips, D., & Howes, C. (2014, November). *Worthy Work, STILL Unlivable Wages: The early childhood workforce 25 years after the national child care staffing study.* Center for the Study of Child Care Employment, University of California, Berkeley. https://cscce.berkeley.edu/worthy-work-still-unlivable-wages/

Zellman, G. L., Perlman, M., Le, V.-N., & Setodji, C. M. (2008). *Assessing the validity of the Qualistar early learning quality rating and improvement system as a tool for improving child-care quality.* RAND Corporation. https://www.rand.org/pubs/monographs/MG650.html

Zigler, E., Gilliam, W. S., & Barnett, W. S. (2011). *The pre-k debates: Current controversies & issues.* Paul. H. Brookes Publishing.

Author Biography

Deborah Stipek, Ph.D., is the Judy Koch Professor of Education, the Peter E. Haas Faculty Director of the Haas Center for Public Service at Stanford University, and has served as Dean of the Stanford Graduate School of Education. Her scholarship focuses on early childhood education. She is particularly concerned about policies and practices that afford children of color and children living in poverty the educational advantages of their more affluent peers. Her current focus is on strategies to develop young children's basic academic skills while supporting their social-emotional development and motivation. In addition to her scholarship, she was a Society for Research in Child Development Fellow, working in the office of Senator Bill Bradley; she served for 5 years on the Board on Children, Youth, and Families of the National Academy of Sciences; and she is a member of the National Academy of Education. She currently chairs the Heising-Simons Development and Research on Early Math Education Network and is a member of California's Master Plan in Early Learning and Care Planning Team.

CALIFORNIA QUALITY RATING AND IMPROVEMENT SYSTEM (CA-QRIS)
QUALITY CONTINUUM FRAMEWORK –RATING MATRIX WITH ELEMENTS AND POINTS FOR CONSORTIA COMMON TIERS 1, 3, AND 4

ELEMENT	1 POINT	2 POINTS	3 POINTS	4 POINTS	5 POINTS
CORE I: CHILD DEVELOPMENT AND SCHOOL READINESS					
1. Child Observation	☐ Not required	☐ Program uses evidence-based child assessment/observation tool annually that covers all five domains of development	☐ Program uses valid and reliable child assessment/ observation tool aligned with CA *Foundations & Frameworks*[1] twice a year	☐ DRDP (minimum twice a year) and results used to inform curriculum planning	☐ Program uses DRDP twice a year and uploads into DRDP Tech and results used to inform curriculum planning
2. Developmental and Health Screenings	☐ Meets Title 22 Regulations	☐ Health Screening Form (Community Care *Licensing form LIC 701 "Physician's Report - Child Care Centers" or equivalent)* used at entry, then: **1.** Annually **OR** **2.** Ensures vision and hearing screenings are conducted annually	☐ Program works with families to ensure screening of all children using a **valid and reliable developmental screening tool** at entry and as indicated by results thereafter **AND** ☐ Meets Criteria from point level 2	☐ Program works with families to ensure screening of all children using the **ASQ** at entry and as indicated by results thereafter **AND** ☐ Meets Criteria from point level 2	☐ Program works with families to ensure screening of all children using the **ASQ & ASQ-SE**, if indicated, at entry, then as indicated by results thereafter **AND** ☐ Program staff uses children's screening results to make referrals and implement intervention strategies and adaptations as appropriate **AND** ☐ Meets Criteria from point level 2
CORE II: TEACHERS AND TEACHING					
3. Minimum Qualifications for Lead Teacher/ Family Child Care Home (FCCH)	☐ Meets Title 22 Regulations [**Center:** 12 units of Early Childhood Education (ECE)/Child Development (CD) **FCCH:** 15 hours of training on preventive health practices]	☐ **Center:** 24 units of ECE/CD[2] **OR** Associate Teacher Permit ☐ **FCCH:** 12 units of ECE/CD **OR** Associate Teacher Permit	☐ 24 units of ECE/CD + 16 units of General Education **OR** Teacher Permit **AND** ☐ 21 hours professional development (PD) annually	☐ Associate's degree (AA/AS) in ECE/CD (or closely related field) **OR** AA/AS in any field plus 24 units of ECE/CD **OR** Site Supervisor Permit **AND** ☐ 21 hours PD annually	☐ Bachelor's degree in ECE/CD (or closely related field) **OR** BA/BS in any field plus/with 24 units of ECE/CD (or master's degree in ECE/CD) **OR** Program Director Permit **AND** ☐ 21 hours PD annually
4. Effective Teacher–Child Interactions: CLASS Assessments (*Use tool for appropriate age group as available)	☐ Not Required	☐ Familiarity with CLASS for appropriate age group as available by one representative from the site	☐ Independent CLASS assessment by reliable observer to inform the program's professional development/improvement plan	☐ Independent CLASS assessment by reliable observer with minimum CLASS scores: **Pre-K** ▪ Emotional Support – 5 ▪ Instructional Support –3 ▪ Classroom Organization – 5 **Toddler** ▪ Emotional & Behavioral Support – 5	☐ Independent assessment with CLASS with minimum CLASS scores: **Pre-K** ▪ Emotional Support – 5.5 ▪ Instructional Support – 3.5 ▪ Classroom Organization – 5.5 **Toddler** ▪ Emotional & Behavioral Support – 5.5 ▪ Engaged Support for Learning – 4 **Infant** ▪ Responsive Caregiving (RC) – 5.5

1. Approved assessments are: Creative Curriculum GOLD, Early Learning Scale by National Institute of Early Education Research (NIEER), and Brigance Inventory of Early Development III.
2. For all ECE/CD units, the core eight are desired but not required.

Note: Point values are not indicative of Tiers 1–5 but reflect a range of points that can be earned toward assigning a tier rating (see Total Point Range).

ELEMENT	1 POINT	2 POINTS	3 POINTS	4 POINTS	5 POINTS
				▪ Engaged Support for Learning – 3.5 **Infant** ▪ Responsive Caregiving (RC) – 5.0	
CORE III: PROGRAM AND ENVIRONMENT - Administration and Leadership					
5. Ratios and Group Size (Centers Only beyond licensing regulations)	☐ **Center:** Title 22 Regulations **Infant** Ratio of 1:4 **Toddler Option** Ratio of 1:6 **Preschool** Ratio of 1:12 ☐ **FCCH:** Title 22 Regulations *(excluded from point values in ratio and group size)*	☐ **Center - Ratio: Group Size** **Infant/Toddler** – 4:16 **Toddler** – 3:18 **Preschool** – 3:36	☐ **Center - Ratio: Group Size** **Infant/Toddler**– 3:12 **Toddler** – 2:12 **Preschool**– 2:24	☐ **Center - Ratio: Group Size** **Infant/Toddler** – 3:12 or 2:8 **Toddler** – 2:10 **Preschool** – 3:24 or 2:20	☐ **Center - Ratio: Group Size** **Infant/Toddler** – 3:9 or better **Toddler** – 3:12 or better **Preschool** – 1:8 ratio and group size of no more than 20
6. Program Environment Rating Scale(s) (Use tool for appropriate setting: ECERS-R, ITERS-R, FCCERS-R)	☐ Not Required	☐ Familiarity with ERS and every classroom uses ERS as a part of a Quality Improvement Plan	☐ 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
7. Director Qualifications (Centers Only)	☐ 12 units ECE/CD+ 3 units management/ administration	☐ 24 units ECE/CD + 16 units General Education +/with 3 units management/ administration **OR** Master Teacher Permit	☐ Associate's degree with 24 units ECE/CD +/with 6 units management/ administration and 2 units supervision **OR** Site Supervisor Permit **AND** ☐ 21 hours PD annually	☐ Bachelor's degree with 24 units ECE/CD +/with 8 units management/ administration **OR** Program Director Permit **AND** ☐ 21 hours PD annually	☐ Master's degree with 30 units ECE/CD including specialized courses +/with 8 units management/ administration, **OR** Administrative Credential **AND** ☐ 21 hours PD annually
TOTAL POINT RANGES					
Program Type	**Common-Tier 1**	**Local-Tier 2[3]**	**Common-Tier 3**	**Common-Tier 4**	**Local-Tier 5[4]**
Centers 7 Elements for 35 points	**Blocked** (7 points) – Must Meet All Elements	**Point Range** 8 to 19	**Point Range** 20 to 25	**Point Range** 26 to 31	**Point Range** 32 and above
FCCHs 5 Elements for 25 points	**Blocked** (5 points) – Must Meet All Elements	**Point Range** 6 to 13	**Point Range** 14 to 17	**Point Range** 18 to 21	**Point Range** 22 and above

3. Local-Tier 2: Local decision if Blocked or Points and if there are additional elements.

4. Local-Tier 5: Local decision if there are additional elements included California Department of Education, February 2014 updated on May 28, 2015; effective July 1, 2015.

Policy Analysis for California Education (PACE)

Improving education policy and practice and advancing equity through evidence

PACE is an independent, non-partisan research center led by faculty directors at Stanford University, the University of Southern California, the University of California Davis, the University of California Los Angeles, and the University of California Berkeley. Founded in 1983, PACE bridges the gap between research, policy, and practice, working with scholars from California's leading universities and with state and local decision makers to achieve improvement in performance and more equitable outcomes at all levels of California's education system, from early childhood to postsecondary education and training. We do this through:

1. bringing evidence to bear on the most critical issues facing our state;
2. making research evidence accessible; and
3. leveraging partnership and collaboration to drive system improvement.

PACE
Policy Analysis for California Education

Stanford Graduate School of Education
520 Galvez Mall, CERAS 401
Stanford, CA 94305-3001
Phone: (650) 724-2832
Fax: (650) 723-9931

edpolicyinca.org