



Quality Rated Validation Study Report #1

What Makes Up a Quality Rated Star Rating? An In-Depth Look at the Criteria, Standards, and Components

Diane M. Early, Kelly L. Maxwell, Nadia S. Orfali, and Weilin Li

Suggested citation

Early, D. M., Maxwell, K. L., Orfali, N. S., & Li, W. (2017). *Quality Rated Validation Study Report #1 What Makes Up a Quality Rated Star Rating? An In-Depth Look at the Criteria, Standards, and Components*. Chapel Hill, NC: Child Trends.

Acknowledgments

We would like to thank Kristin Bernhard, Catherine Broussard, Jessie Bruno, Meghan Dean, Randy Hudgins, Denise Jensen, Rob O’Callaghan, Nnenna Ogbu, Bentley Ponder, and Pam Stevens from the Georgia Department of Early Care and Learning for their support of this work. We would also like to thank Nicole Forry, our colleague at Child Trends, for her review.

This project was funded by the U.S. Department of Education and U.S. Department of Health and Human Services through Georgia’s Race to the Top-Early Learning Challenge grant, and we would like to thank the Early Learning Challenge grant leaders for their guidance and support of this evaluation. We hope that this and future Quality Rated validation reports will be used to support the continued improvement of Quality Rated and the programs that serve young children in Georgia.

Copyright Child Trends 2017 | Publication #2017-27



Contents

Report Highlights	1
The Quality Rated Validation Approach	2
Overview of Quality Rated Validation Reports.....	2
Current Report: Purpose and Process	3
Quality Rated Star Levels	3
Methods.....	6
Findings.....	6
1. What is the distribution of star ratings?.....	6
2. On which components are programs earning points?.....	7
3. How do ratings compare for programs that do and do not include Georgia's Pre-K or Head Start?	8
4. How is the final star rating determined?.....	10
Conclusions and Recommendations	12
References	14
Appendices	15
Appendix A: What types of programs are included in Others?	15
Appendix B: How do the component scores relate to the star ratings?.....	16
Appendix C: How do programs score on the Process Quality component (i.e., ERS)?..	17
Appendix D: What is the distribution of scores for each Quality Rated standard in the portfolio?.....	19
Appendix E: What is the distribution of points earned in the individual criteria that make up the Quality Rated standards in the portfolio?.....	20
Appendix F: How do Child Care Learning Centers with Georgia's Pre-K earn points in Quality Rated?.....	21
Appendix G: How do Child Care Learning Centers and Others with Head Start funding earn points in Quality Rated?.....	23
Appendix H: How do the various Quality Rated criteria, standards, and ERS relate to one another?	25
Appendix I: Where are 0-star programs typically losing points?	26

Report Highlights

Quality Rated is Georgia's systemic approach to assessing, improving, and communicating the level of quality in early childhood and school-age care and education programs. Programs are assigned a star rating based on a combination of a portfolio designed to measure Structural Quality and observations to measure Process Quality.

The portfolio measures Structural Quality using five standards: *Director and Teacher Qualifications; Child Health, Nutrition, and Physical Activity; Family Engagement; Intentional Teaching Practices; and Teacher:Student Ratios.* Each standard is measured using multiple criteria or indicators. Process Quality observations are conducted using the Environment Rating Scales (ERS), a widely used group of tools designed to measure quality in early childhood settings by observing activities, interactions, and health and safety practices.

This report explores data from the 1,516 programs that were rated as of May 31, 2017. This includes 1,034 Child Care Learning Centers (CCLCs), 402 Family Child Care Learning Homes (FCCLHs), and 80 programs categorized as *Other*.

Findings include:

- 1. The most common rating was 2-star, followed by 1-star, and there were some differences in the distribution of star ratings among different types of programs.** Across all types of settings, over three quarters of programs were either 1- or 2-star. A higher proportion of FCCLH programs than CCLC programs attained the highest rating (3-star) ([see Figure 2](#)).
- 2. Programs earned a higher proportion of the available Structural Quality points than Process Quality points.** This was true for all types of programs. CCLCs attained fewer Process Quality points than either FCCLHs or Others, on average ([see Figure 3](#)).
- 3. Programs that are held to more rigorous standards generally attained a higher star rating.** Both Georgia's Pre-K and Head Start have program standards and monitoring requirements that are more stringent than child care licensing standards. Programs with these types of funding were more likely to be rated at the upper star levels than those without ([see Figure 4](#) and [Figure 5](#)).
- 4. The star rating is driven almost entirely by the Process Quality component (i.e., ERS score).** Ninety-four percent of the rated programs would have received the same rating using only the points from the Process Quality component. Structural Quality points played a role in determining the rating of only a small minority of programs ([see Table 2](#) and [Table 3](#)).

The Quality Rated Validation Approach

Quality Rated—Georgia’s Tiered Quality Rating and Improvement System (TQRIS)—is Georgia’s systematic approach to assessing, improving, and communicating the level of quality in early childhood and school-age care and education programs.

As part of Georgia’s Race to the Top Early Learning Challenge (RTT-ELC) grant application, Georgia’s Department of Early Care and Learning (DECAL) committed to expanding and evaluating Quality Rated. The star rating is one component of the Quality Rated system; other components include supports to programs to improve quality, and communication with the public about quality. Using RTT-ELC funding, DECAL has contracted with Child Trends, in partnership with Georgia State University, to validate the Quality Rated rating system.

Zellman and Fiene (2012) define TQRIS validation as a “*multi-step process* that assesses the degree to which design decisions about program quality standards and measurement strategies are resulting in accurate and meaningful ratings.” TQRIS validation is not a one-time event, but rather a series of activities that support the continuous improvement of the TQRIS. Validation studies do not yield a *yes* or *no* answer (i.e., yes, the TQRIS is valid) but instead provide detailed information used to support the continuous improvement and refinement of the TQRIS.

The overarching goal of the Quality Rated Validation Project is to provide Georgia’s early childhood leaders with high-quality information about the rating system that can be used to strengthen it. Within that goal, there are three primary objectives:

1. Quality Rated leaders will have a better understanding of the strengths and limitations of the Quality Rated administrative data system, as well as an understanding of how the rating system is functioning (e.g., associations among various components of the rating).
2. Quality Rated leaders will have a better understanding of the extent to which Quality Rated ratings are accurate and meaningful indicators of program quality.
3. Quality Rated leaders will have a better understanding of the extent to which Quality Rated program ratings, as well as the standards that comprise the ratings, are related to children’s development and learning.

This report, which is part of the Quality Rated Validation Project, presents administrative data as a first step in validating the rating system.

Overview of Quality Rated Validation Reports

A series of four reports will be developed as part of the Quality Rated Validation Project. This report is the first in the series and provides information about how the components are working together to create the rating and how ratings differ by program type. Future reports, to be delivered throughout 2017 and 2018, will include additional analysis of administrative data, as well as data that are currently being collected. Child Trends and Georgia State University are collaborating to collect and analyze information from teachers, providers, and directors about their experiences with Quality Rated; independent classroom and program observations; audio recordings of teacher and provider interactions with children to understand language use; and assessments of children’s emerging academic and social skills. Combined, this series of reports will address the three primary objectives delineated above and provide data to guide Quality Rated leaders in considering future implementation and revision.

Current Report: Purpose and Process

This report addresses the first of the three primary objectives by analyzing Quality Rated administrative data to understand how the criteria, standards, and components are working together to create the ratings. The report focuses on the function of the ratings themselves and answers questions such as: Does the distribution of Quality Rated ratings vary by program type or funding source? And, to what extent do the various standards contribute to the final rating?

The questions addressed in the report stem from an analysis plan that was jointly developed by the Child Trends and DECAL research teams, with input from the Quality Rated External Validation Committee. Working closely together, these teams generated research questions that could use existing administrative data to provide insight into the functioning of the rating system. The questions and frequency of analysis are delineated in an analysis plan that is reviewed and revised as needed. By creating this research plan in advance, the teams sought to thoughtfully narrow the scope of the administrative analyses to questions that could be answered on a regular basis and would provide meaningful feedback to DECAL.

Quality Rated Star Levels

One important aspect of Quality Rated is assigning star ratings. The rating process is summarized below. For a detailed description, see the [Quality Rated Program Manual](#).^a

Quality Rated star levels range from 1 to 3. For purposes of this report, programs that complete the rating process but do not meet the criteria for 1, 2, or 3 stars are referred to as *O-star*. From a policy standpoint, DECAL considers these programs to be not rated, and DECAL does not use the term *O-star*. However, because these programs sought a rating, took part in all aspects of the rating process, and have usable data in the Quality Rated system, we believed including them here was important for understanding how the rating is working.

To begin the process of earning a Quality Rated star level, a program director, FCCLH provider, or designee of an eligible program completes a short application online. Once the application is submitted and has been accepted, programs can access Quality Rated resources, take part in an orientation of Quality Rated and the Environment Rating Scales (ERS),^b begin the submission of their portfolio documentation, and receive free technical assistance from their local Child Care Resource and Referral agency (CCR&R) to help them with the rating process. Programs are considered to be *participating* in Quality Rated upon completing the application.

Programs submit evidence in an online portfolio to earn points based on increasingly difficult criteria aligned with five standards. The standards are described here, and the criteria that make up the standards are described in [Table 1](#).

- Standard 1 - *Director and Teacher Qualifications and Professional Development Registry Verification* focuses on the qualifications of staff and continued development of their professional skills.
- Standard 2 - *Child Health, Nutrition, and Physical Activity* focuses on health practices that go above and beyond licensing standards.
- Standard 3 - *Family Engagement* focuses on practices that connect families and providers with communities to improve child outcomes.
- Standard 4 - *Intentional Teaching Practices* focuses on alignment of teaching practices and curriculum to Georgia's Early Learning and Development Standards (GELDS), and on integrating planning and assessment.
- Standard 5 - *Teacher:Student Ratio Requirements* focuses on having smaller groups of children per adult to enhance the learning environment.

^a https://qualityrated.dec.al.ga.gov/Content/Documents/PM_ProgramManual.pdf

^b ERS is a group of observational tools used to assess process quality in early childhood care and education programs. See <http://ersi.info/> for more information.

Once the online portfolio has been accepted, a Quality Rated assessor from DECAL conducts an unannounced ERS observation. If there is more than one classroom in the program, observations are conducted in one third of the classrooms of each age group, which are selected at random. The assessor uses the version of the ERS that is appropriate for the setting or the age group. If multiple ERS observations are conducted, the average ERS score is used in the rating process; however, if any classroom within a program scores less than a 3.0 on the ERS, the program cannot earn more than a 1-star rating.

The score from the online portfolio (range: 1-104 for all programs except for Family Child Care Learning Homes with no additional staff members, for which the range is 1-86) is converted to Structural Quality points (range: 0-15), and the average ERS score (range: 1-7) is converted to Process Quality points (range: 0-30). Programs can also earn Bonus Points (range: 0-4) based on accreditations such as National Association for the Education of Young Children. [Figure 1](#) shows the mapping of criteria onto standards and the conversion to the final rating components.

Total points—calculated by adding together Structural Quality points, Process Quality points, and Bonus points—determines the program’s star level, where 0-14 total points means a program earns no (0) stars, 15-24 total points is a 1-star, 25-35 total points is a 2-star, and 36-45 total points is a 3-star.

Table 1. Evidence programs submit and how programs earn points for each criterion

Standard	Criteria	How programs earn points
1	1.1-Director/Provider Education	Career Levels in the Georgia Professional Development System for Early Childhood Educators (GaPDS)
1	1.2-Director/Provider PLP (Professional Learning Plan)	Copy of PLP addressing administrative competencies and training hour goals
1	1.3-Director/Provider Annual Training	Hours of training in GaPDS
1	1.4-Teacher Education	Percentages of lead and assistant teachers meeting specified career levels in GaPDS
1	1.5-Teacher PLP	Copy of PLP addressing early childhood education competencies and goal hours of training
1	1.6-Teacher Annual Training	Percentages of lead and assistant teachers meeting hours of training in GaPDS
2	2.1-Nutrition and Physical Activity Plan	Completes self-assessment and writes an improvement plan addressing goals
2	2.2-First Aid/CPR Certification	Percentages of staff with current certification
2	2.3-Health Information/Family Resources	Submits samples of written materials and policies related to early interventions of health issues
3	3.1-Five Protective Factors	Percentages of staff who have completed the Strengthening Families training (verified in GaPDS)
3	3.2-Family Engagement Self-Assessment	Completes self-assessment and writes an improvement plan addressing goals
3	3.3-Family Engagement Written Evidence	Submits samples of written policies or evidence of family activities or resources that strengthen family connections

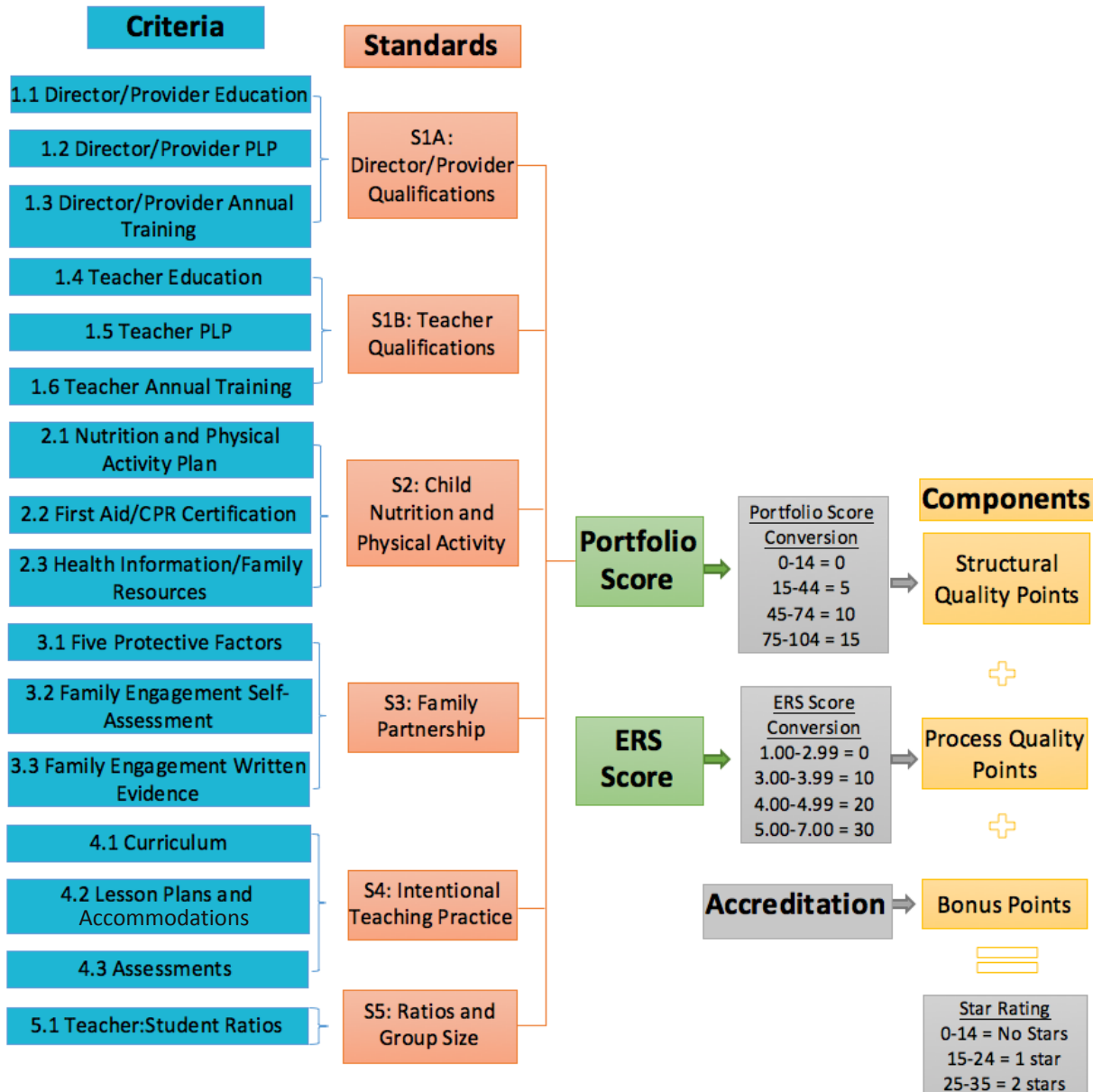
Table 1. Evidence programs submit and how programs earn points for each criterion (cont.)

Standard	Criteria	How programs earn points
4	4.1-Curriculum	Submits a description of how the curriculum supports the five learning domains; is culturally, linguistically, and developmentally appropriate; and is aligned with GELDS
4	4.2-Lesson Plans and Accommodations for Every Child	Submits one week of lesson plans per age group
4	4.3-Assessments	Submits a narrative describing use of assessments
5	5.1-Teacher:Student Ratios	Submits counts of teachers and students for one day for each classroom, within 30 calendar days of portfolio submission

Source: Quality Rated Program Manual

Figure 1. Diagram of the Quality Rated system

The portfolio score, made of up of criteria organized into standards, is converted into Structural Quality points; ERS scores are converted into Process Quality points.



Methods

This report relies on data from the Quality Rated Administrative Data System collected and maintained by DECAL as part of the process of assigning a star rating. This system, developed and built internally by DECAL, houses all information used to assign the star rating, including scores on criteria, standards, and components, as well as some descriptive information about rated programs.

The current report includes all 1,516 programs that had completed the rating process (0-, 1-, 2-, and 3-star) as of May 31, 2017. This includes 1,034 Child Care Learning Centers (CCLCs), 402 Family Child Care Learning Homes (FCCLHs), and 80 programs categorized as *Others*. In Georgia, Group Day Care Homes and unlicensed programs that are subject to different government oversight (e.g., programs on military bases, some Head Start programs, some programs housed at universities or colleges, and some school-based programs^c) can elect to take part in Quality Rated. DECAL calls this category *Others*, and we use that label throughout the report. See [Appendix A](#) for a breakdown of the programs included in this category.

For this report, data from the most recent rating were used for programs that had been rated more than once.^d These 1,516 programs represent 30 percent of all Quality Rated eligible programs statewide (33 percent of CCLCs, 23 percent of FCCLHs, and 58 percent of Others).

Findings

1. What is the distribution of star ratings?

As seen in [Figure 2](#), 2-star programs were the most common, followed by 1-star, and over three quarters of programs fell into one of these two categories. The average star rating was 1.64 across all programs.

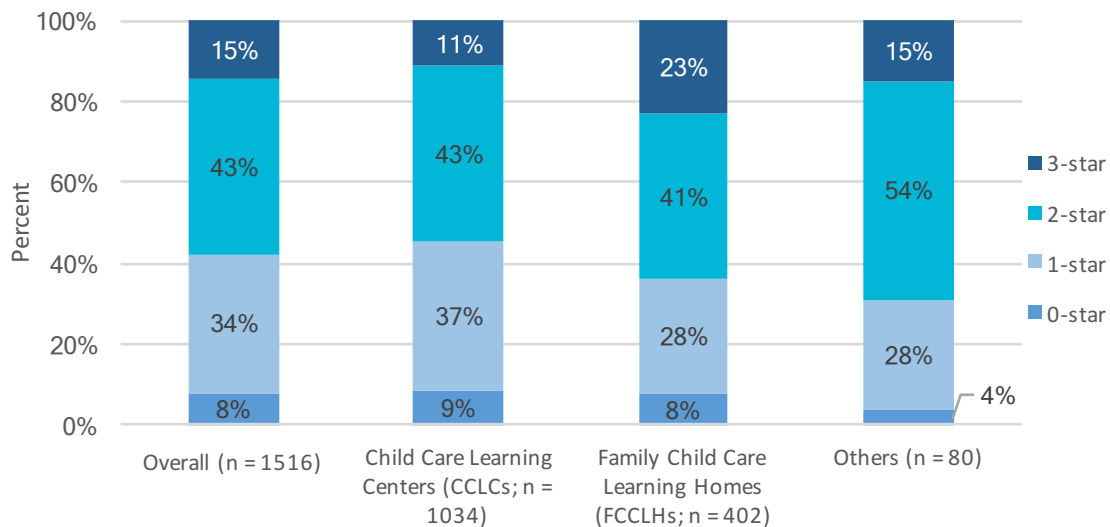
Comparisons among program types demonstrate which aspects of the rating are more or less challenging for programs that vary in licensing requirements and may serve different populations. The percentage rated at the top level (3-star) was significantly higher and the percentage rated at the 1-star level was significantly lower for FCCLHs compared to CCLCs. In addition, there were significantly more 0-star CCLCs than 0-star Others and significantly more 2-star Others than 2-star FCCLHs. CCLCs had a significantly lower average star rating (1.57) than FCCLH programs (1.79) and Others (1.80).

^c Typically, local school systems are not eligible to take part in Quality Rated. The local school systems included in the Others group in the current analyses took part in a Quality Rated pilot initiative to determine the feasibility of including local school systems in Quality Rated in the future.

^d Quality Rated star ratings are awarded for a 3-year period with an annual renewal requirement. At their annual renewal, programs can choose to maintain their rating, complete a shortened portfolio to apply for a Continuous Quality Improvement grant, or request a reassessment (if they meet certain requirements for each option, such as a minimum portfolio score or current star rating). Programs that request a reassessment submit a new portfolio and have an ERS observation, just like their initial rating process. Out of the 1,516 programs in our sample, 263 had gone through a reassessment.

Figure 2. Distribution of star ratings

Over three-quarters of all programs are 1- or 2-star.



Source: DECAL's administrative data as of May 31st, 2017

2. On which components are programs earning points?

Programs earn points in three components (Structural Quality, Process Quality, and Bonus) that are added together to determine their final rating. To better understand the final scores and areas of relative difficulty, we examined the components individually. [Figure 3](#) presents the average number of points that different types of programs earned for Structural Quality and Process Quality. The top of the striped areas indicates the maximum number of points available.

All three types of programs attained a high proportion of the available Structural Quality points, which are derived from the portfolio that providers submit. More points are available in Process Quality, which is derived from the ERS, but programs attained a smaller proportion of the Process Quality points overall. Bonus points are awarded for having a national accreditation, but programs generally earned very few Bonus points (not pictured). On average, out of the four possible Bonus points, CCLCs earned 0.30, FCCLHs earned 0.16, and Others earned 0.38.^e

CCLCs earned significantly fewer Structural Quality points than Others, but there were no additional significant differences among program types. For Process Quality, on average, CCLCs attained significantly fewer points than either FCCLHs or Others, but there were no differences between FCCLHs and Others.

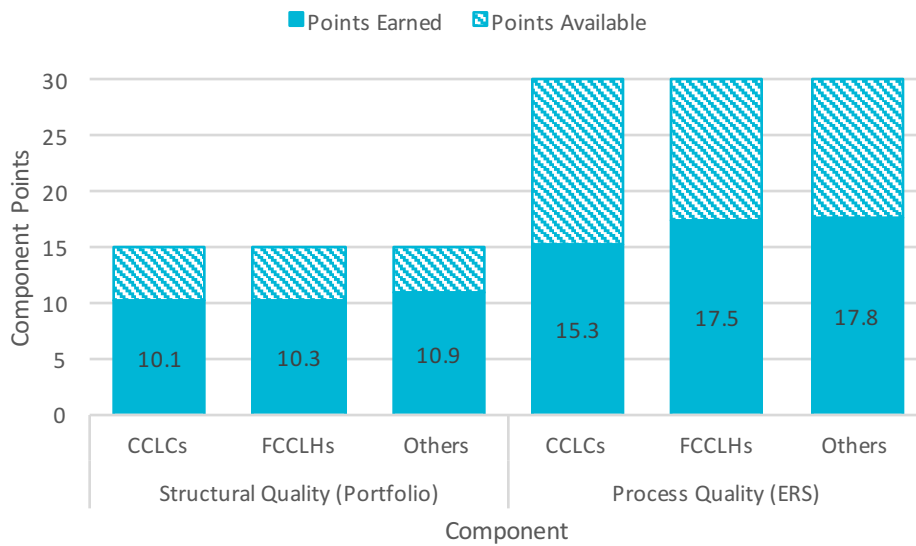
Across all programs, 22 percent earned all 15 points available for Structural Quality and 15 percent earned all 30 points available for Process Quality. Programs of different types were equally likely to earn all available Structural Quality points (22 percent of CCLCs, 25 percent of FCCLHs, and 23 percent of Others). However, a smaller portion (12 percent) of CCLCs earned all available Process Quality points, compared to 25 percent of FCCLHs and 15 percent of Others.

For more information about how many points programs earn in each component, broken down by star-level, see [Appendix B](#). For additional information about how programs score on the Process Quality component (i.e., ERS), see [Appendix C](#). For a breakdown of scores for each Quality Rated standard and criteria, see [Appendix D](#) and [Appendix E](#), respectively.

^e The remainder of this report does not include details regarding Bonus points because few programs earn them and they play a very small role in the rating.

Figure 3. Average component points earned

Programs earned a high proportion of the available Structural Quality points; they earned a smaller proportion of the available Process Quality points



Source: DECAL's administrative data as of May 31st, 2017

3. How do ratings compare for programs that do and do not include Georgia's Pre-K or Head Start?

Some Quality Rated programs receive funding from Georgia's Pre-K and/or Head Start. These additional funds, and their accompanying standards and monitoring, may lead to differences in their star rating.

Georgia's Pre-K

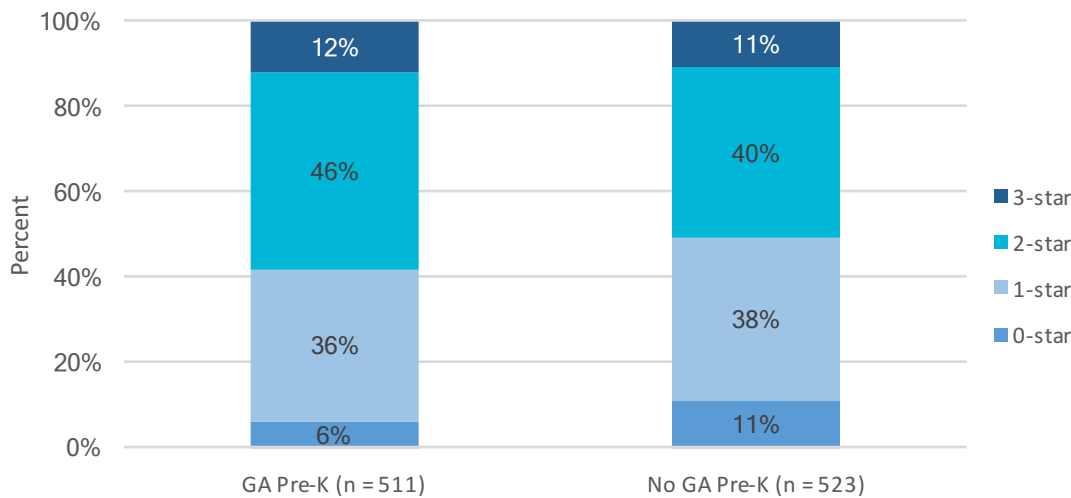
Georgia's Pre-K Program is a 25-year-old lottery-funded educational program for Georgia's 4-year-olds to prepare them for kindergarten. Georgia's Pre-K classrooms serve children the year before kindergarten at no cost to the family. Just over half (55 percent) of all Georgia's Pre-K classrooms are in a CLCC and the remainder (45 percent) are in a public school, but only those in a CCLC are typically eligible to participate in Quality Rated. Classrooms funded by Georgia's Pre-K in CCLCs must meet higher standards than those required for licensing, and the director and pre-K teachers receive additional supports from the state. The classrooms and teachers at the CCLC that are not supported by Georgia's Pre-K do not have to meet the higher standards, and CCLCs that include Georgia's Pre-K are not required to take part in Quality Rated. It is possible, however, that the additional standards, supports, and resources provided by Georgia's Pre-K have indirect benefits that raise the general level of quality at the CCLC, or that higher quality CCLCs elect to participate in Georgia's Pre-K. For those reasons, we anticipated that programs that took part in Georgia's Pre-K would generally have higher star ratings.

About half (49 percent) of the Quality Rated CLCCs had one or more Georgia's Pre-K classrooms. For purposes of Quality Rated, Georgia's Pre-K classrooms are treated like other center-based classrooms serving preschool-aged children. ERS observers randomly select preschool classrooms for observation, and all Georgia's Pre-K classrooms in the center are included in the pool for possible random selection.

CCLCs that included Georgia’s Pre-K were significantly more likely to be rated as 2-star and less likely to be rated as 0- star than those that did not have Georgia’s Pre-K (see [Figure 4](#)). For more information about how CCLCs that include Georgia’s Pre-K scored on specific Quality Rated components, see [Appendix F](#).

Figure 4. Star ratings for programs with and without Georgia’s Pre-K (CCLCs only)

Programs with Georgia’s Pre-K were more likely to be 2- or 3-star.



Source: DECAL’s administrative data as of May 31st, 2017

Head Start funding

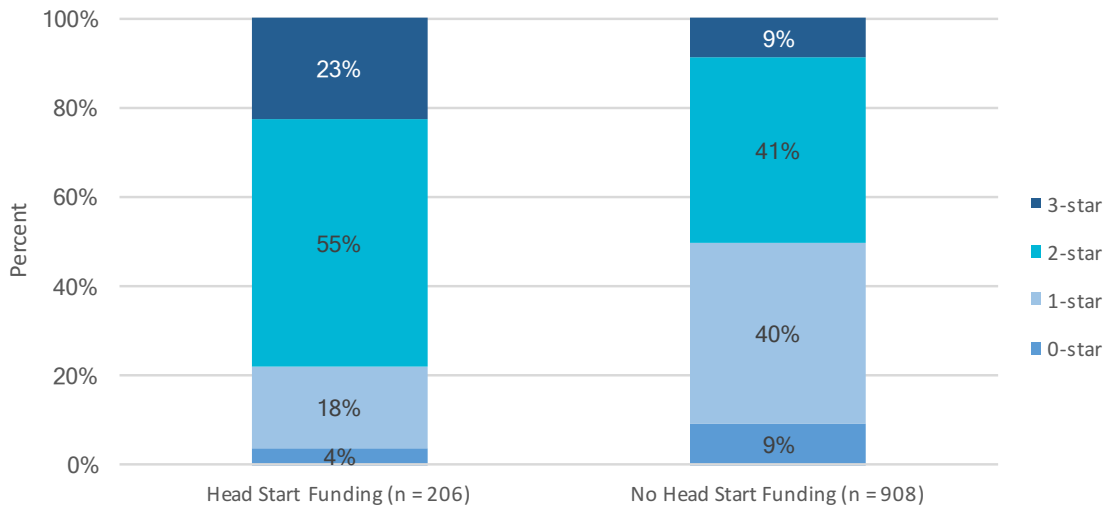
Among CCLCs and Others in Quality Rated, about 18 percent received Head Start or Early Head Start funds.^f These programs may have a variety of different configurations, including traditional Head Start or Early Head Start programs, traditional center-based programs that operate one or more Head Start classrooms, traditional center-based programs that receive Early Head Start Child Care Partnership funds, and center-based programs that use Head Start funds to cover the costs of individual children in preschool classrooms. Regardless, we anticipated that the presence of the Head Start funding—coupled with the addition of Head Start regulations and monitoring—would be associated with higher star ratings among programs with Head Start funding.

As seen in [Figure 5](#), CCLCs and Others with Head Start funding were significantly more likely to be rated at the 2- and 3- star level, and less likely to be rated at the 0- and 1-star level than CCLCs and Others without Head Start. For more information about how CCLCs and Others with and without Head Start funding scored on Quality Rated components, see [Appendix G](#).

^f To determine which programs had Head Start or Early Head Start funds, DECAL used the *Head Start Centers Locations Datasets* as of December 2016 (<http://eclkc.ohs.acf.hhs.gov/hslc/data/center-data>), unduplicated by site and grantee. DECAL matched those data to Quality Rated and licensing data from March 2017, but some judgement had to be used when programs had similar, but not identical, names and addresses. This is a manual process that occurs once a year; therefore, some differences in licensing status may occur between March 2017 and the Quality Rated administrative dataset used in this report. In addition, Others with a provider number prefix that designated Head Start or Early Head Start (i.e., GAHS or GAEHS) were included in the Head Start sample. Note that when programs have Head Start funds, it is not possible to know how integrated the Head Start program is with other types of programs (e.g., licensed, private, or public child care, Georgia’s Pre-K) that are in the same physical location.

Figure 5. Star ratings for programs with and without Head Start funding (CCLCs and Others Only)

Child Care Learning Centers and Others with Head Start funding were more likely to be rated at the 2- and 3-star level than those without Head Start funding.



Source: DECAL's administrative data as of May 31st, 2017

4. How is the final star rating determined?

In creating a TQRIS, each state not only determines what criteria, standards, and observational tools to include, but also how to combine those elements to create a final score. Due to the complex nature of these systems and the need to aggregate many different types of information into a single value, it is not always obvious to what extent each element drives the rating. This section explores how various parts of the Quality Rated system contribute to the rating.

To what extent do the individual standards and quality points contribute to the final rating?

The first six rows of [Table 2](#) show how strongly related each of the standards is to the final star rating. The final two rows show how strongly related the Structural Quality points (calculated by transforming the total portfolio score into a new scale) and the Process Quality points (calculated by transforming the ERS score into the same scale as the Structural Quality points) are to the final star rating. The values on this table are correlation coefficients (*r*).⁹ A correlation of .30 is considered weak, a correlation of .50 is moderate, and a correlation of .70 is strong (Evans, 1996).

Across all programs, each standard was significantly associated with the star rating, but the associations were weak (see [Table 2](#)). Within program type, most standards were significantly related to the final star rating. The only exceptions were standards 2 through 5 for Others, which were not significantly related to the final star rating.

Structural Quality points and Process Quality points were significantly related to the final star rating for all programs and for each program type. The striking part of this table is the fact that the Process Quality points were correlated almost perfectly with the star rating. Put another way, once the ERS score is determined, the star rating is mostly determined. Structural Quality, derived from the portfolio, is contributing little to the final rating.

⁹ Correlation coefficients can range from -1.0, (indicating that the two are perfectly associated, but when one is higher the other is lower), to 0 (no association at all), to 1.0 (perfect association with both being higher or both lower).

Table 2. Correlations between standards, components, and star ratings*Star ratings are determined almost entirely by Process Quality points (ERS Scores).*

Standards and components	Correlations with star ratings			
	All programs	CCLCs	FCCLHs	Others
S1A: Director Qualifications	0.30	0.29	0.35	0.27
S1B: Teacher Qualifications	0.13	0.25	N/A	0.25
S2: Nutrition & Activity	0.20	0.18	0.24	0.13
S3: Family Partnership	0.22	0.21	0.25	0.15
S4: Intentional Teaching	0.22	0.20	0.33	0.06
S5: Ratios and Group Size	0.13	0.16	0.14	-0.03
Structural Quality Points (Portfolio)	0.42	0.41	0.45	0.34
Process Quality Points (ERS)	0.96	0.95	0.96	0.98

Source: DECAL's administrative data as of May 31st, 2017**How closely are ERS and star ratings aligned?**

To further explore the association between ERS scores and star ratings, we calculated the star rating each program would have received if its average ERS score alone had been used to determine the rating, based on the current ERS cut-offs for Process Quality points (i.e., 1.00-2.99 = 0-star; 3.00-3.99 = 1-star; 4.00-4.99 = 2-star; 5.00-7.00 = 3-star). We then compared this ERS-only rating to the program's actual rating as determined by the combination of their Structural Quality, Process Quality, and Bonus points. As seen in [Table 3](#), almost all programs received the same rating when *only* average ERS scores were used.

Of the 1,513 programs with ERS data available, 1,422 (94 percent) would have received the same rating if only average ERS score had been used (see bold values on the diagonal). Seventy-five programs (5 percent) were helped by the inclusion Structural Quality and Bonus points (see programs below the diagonal), meaning their rating would be lower than it currently is if only their average ERS score were taken into consideration. On the other hand, 16 programs (1 percent) received a lower star rating than they would have if their average ERS score had been the only information used (see programs above the diagonal). Of these 16, five were rated 1-star despite having an average ERS score that corresponds to 2 stars, because at least one of their classrooms scored below a 3.0 on the ERS. (As noted earlier, if any classroom within a program receives an ERS score below 3.0, the program cannot attain more than a 1-star rating.) The remaining 11 programs had very low portfolio scores, which caused their actual star rating to be lower than the rating they would have achieved if only their average ERS score had been used.

Table 3. Ratings predicted by ERS scores alone

Few programs would have received a different rating if the ERS score were the only information considered.

		Star rating predicted by average ERS score alone			
		0-star	1-star	2-star	3-star
Actual star rating	0-star	121	1	0	0
	1-star	11	501	5	0
	2-star	0	64	578	10
	3-star	0	0	0	222

Source: DECAL's administrative data as of May 31st, 2017

Additional analyses are included in the Appendices of this report. [Appendix H](#) includes information about how the various Quality Rated criteria, standards, and ERS relate to one another. [Appendix I](#) shows where 0-star programs typically lose points.

Conclusions and Recommendations

The Quality Rated Administrative Data System contains a wealth of information that can help DECAL better understand how the various pieces of Quality Rated come together to determine the rating. The data system includes extensive detailed information with very little missing data. The administrative data provide clear documentation of how programs attain ratings, serving as a critical support in creating a system that is fair to programs that take part.

The Quality Rated star rating is derived from a combination of Structural Quality, Process Quality, and Bonus points, and this report focused on how the components fit together to create a rating. An important finding from this report was that the star rating is determined almost exclusively by Process Quality points, as measured by the ERS. In fact, Structural Quality and Bonus points only affected the star rating of 6 percent of programs. Put another way, despite the intensive effort on the part of providers and Quality Rated staff to complete and score the portfolio, it had a very small impact on the overall rating.

To decide whether changes need to be made to address this issue, it is important to consider the time and effort on the part of programs and DECAL staff dedicated to each component, as well as the value DECAL places on each component. From a measurement perspective, it is better to rely on multiple sources of high-quality information in determining the rating. Although the ERS is a strong tool, it has limitations such as relying on a single day of observation, requiring a high level of observer inference, and taking only classroom (rather than program) quality into account. Thus, we encourage DECAL to consider how to meaningfully gather and use information in addition to the ERS that can effectively and efficiently distinguish higher quality programs from lower quality programs.

Based on this report and the expertise of the authors, we recommend DECAL consider the following when revising the Quality Rated star-rating system:

- **Balance rating and quality improvement.** We encourage DECAL leaders to consider revisions to the Quality Rated rating from the broad perspective of balancing resources required to develop and implement a valid rating of program quality against the need for resources to support quality improvement. We do not intend to imply that DECAL should heavily invest additional resources in its ratings. Most programs across the country still fall in the low- to-medium range of quality (Burchinal, Magnuson, Powell, & Hong, 2015), so the *I* (improvement) in TQRIS is arguably more important than the *R* (rating).

- **Revise the star rating system so that quality indicators other than the ERS play a meaningful role.** We suggest that Quality Rated leaders revise the star rating to ensure that more information is used. As part of this recommendation, we suggest carefully reviewing the information in the portfolio. Compiling the portfolio is time-consuming for programs, and it is not feasible for DECAL to assess the quality or level of implementation for some criteria. Additionally, we suggest DECAL review findings from the various validation studies in other states to determine which indicators meaningfully distinguish levels of quality and consider how best to weight them in the rating. Child Trends and Georgia State University’s current data collection efforts in all levels of Quality Rated programs will provide some data to guide these decisions.
- **Incorporate perspectives of people receiving, assigning, and using ratings.** Although we hope Quality Rated leaders will use the findings in this report to inform possible revisions to the rating, we also encourage Quality Rated leaders to ask directors, teachers, providers and DECAL staff about their perceptions of the rating process. DECAL has a long history of stakeholder engagement, including a Quality Rated Advisory Committee, made up of providers of all types and at all stages of the Quality Rated process. The group was selected by competitive application and will be an especially strong resource for feedback on any revisions that are considered. Additionally, technical assistance providers may have valuable information to contribute. Talking specifically with Head Start grantees could help Quality Rated leaders consider how best to include Head Start programs, and conversations with centers that are part of Georgia’s Pre-K will help DECAL consider coordination across various monitoring systems and how to efficiently include centers that receive Georgia’s Pre-K funds. Child Trends and Georgia State University’s current data collection effort will also provide some data about providers’ perceptions of the Quality Rated system.
- **Pilot test revisions.** We encourage Quality Rated leaders to pilot test revisions to the rating before fully implementing, to the extent possible. Pilot testing might include gathering some information from a small number of providers or using existing administrative data to test a new way of calculating a rating. Testing revisions before fully implementing them will allow DECAL to ensure that the revisions are working as intended.

To continue supporting quality improvement, we offer the following questions for DECAL’s consideration:

- What technical assistance would be most effective in supporting programs that applied for Quality Rated but were assigned a 0-star rating? This group of providers is interested in quality yet received very low ERS scores. Could targeted technical assistance help these providers improve their quality and earn a higher rating?
- Nearly half of Quality Rated programs (43 percent overall) have earned a 2-star rating. This project’s next report will look at re-ratings and consider what portion of programs that start at 2-star increase to 3-star when re-rated. Are there technical assistance supports that could help the 2-star rated programs move to a 3-star rating at their next renewal?
- The current data system does not clearly identify which programs receive Head Start funds. DECAL has found a way to locate those programs among CCLCs and Others, but it is a labor-intensive process and does not include some Others. Could Quality Rated work with others at DECAL to streamline this process to make it easier and more accurate, possibly by coordinating with the Head Start Collaboration Office?

As a first step in validating Georgia’s TQRIS, this report has presented an overview of Quality Rated’s system for assigning star ratings and information about how Quality Rated’s various components are functioning. Future reports will delve deeper into the administrative data, present new data about users’ perceptions of the system, and consider how Quality Rated is linked to independent measures of quality and children’s development.

References

Burchinal, M., Magnuson, K., Powell, D., & Soliday Hong, S. (2015). Early child care and education. In R. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., Vol. 4, pp. 1-45). Hoboken, NJ: Wiley.

Burchinal, M., Tarullo, L. & Zaslow, M. (2016). *Best practices in creating and adapting Quality Rating and Improvement System (QRIS) rating scales*. OPRE Research Brief #2016-25. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pianta, R.C., La Paro, K.M., & Hamre, B.K. (2008). *Classroom Assessment Scoring System – Pre-K*. Baltimore, MD: Paul H. Brookes Publishing Co., Inc.

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

Appendices

Appendix A: What types of programs are included in Others?

Table A1. Types of programs included in Others

Programs in the Others group include some Head Start, military, and school-based programs.

Program type	Number included in Others
Military (DOD)	6
Early Head Start (GAEHS)	7
Head Start (GAHS)	50
Group Day Care Home (GDCH)	5
Local school system (LSS)*	10
Technical school or college (TEC)	1
College or university (UNIV)	1
Total	80

Source: DECAL's administrative data as of May 31st, 2017

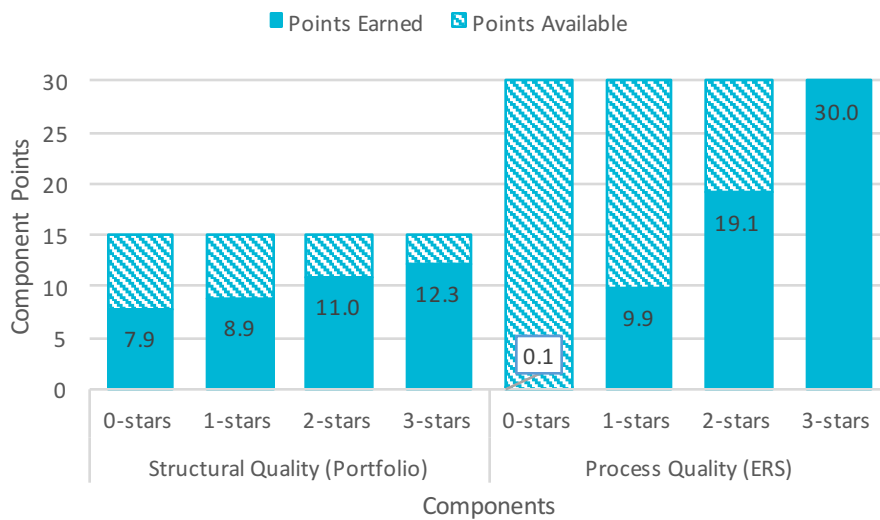
*Typically, local school systems are not eligible to take part in Quality Rated. The local school systems included in the Others group in the current analyses took part in a Quality Rated pilot initiative to determine the feasibility of including LSS in Quality Rated in the future.

Appendix B: How do the component scores relate to the star ratings?

Figure B1 shows the average number of points earned in the Structural Quality and Process Quality components by programs rated at each star level. As would be expected, the points increased with each star level. However, the difference associated with each increase in star level was much greater for Process Quality points than for the Structural Quality points, indicating that Process Quality played a larger role in the star rating.

Figure B1. Average component points earned by programs in each star rating

The difference associated with increases in star rating is stronger for Process Quality than for Structural Quality.



Source: DECAL's administrative data as of May 31st, 2017

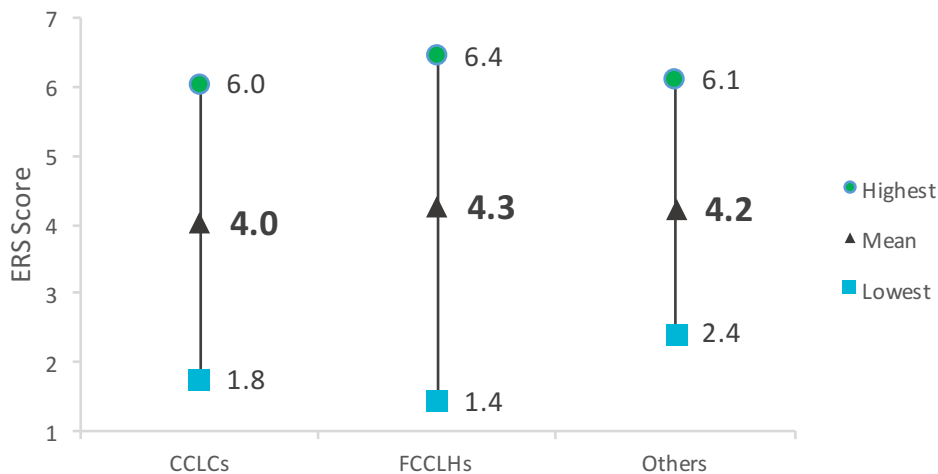
Appendix C: How do programs score on the Process Quality component (i.e., ERS)?

In Quality Rated, Process Quality is measured using the average Environmental Rating Scale (ERS) score. The ERS includes the ECERS-R and ECERS-3,^a which are used in classrooms serving children ages 2 to 5 years old; the ITERS-R, which is designed for use with children from birth to 2.5 years old; and the FCCERS-R, which is designed to be used in family child care programs.^b

As seen on Figure C1, the average ERS score for CCLCs was 4.0, FCCLH programs averaged a 4.3, and the average score for Others was 4.2. The range for all three types of programs was quite large, with some CCLCs and FCCLHs scoring below 2.0 and some in each group scoring a 6.0 or higher. According to the authors of the ERS, a 2.0 is between inadequate and minimal quality; a 6.0 is between good and excellent quality. On average, both FCCLHs and Others scored significantly higher than CCLCs on the ERS. The differences, however, are quite small (0.3 points and 0.2 points, respectively). In the Quality Rated system, an average ERS score of 5.0 or more corresponds to three stars, a score between 4.0 and 4.9 corresponds to two stars, and a score between 3.0 and 3.9 corresponds to one star.

Figure C1. Average ERS scores for each program type

FCCLHs and Others have higher ERS scores than CCLCs.



Source: DECAL's administrative data as of May 31st, 2017

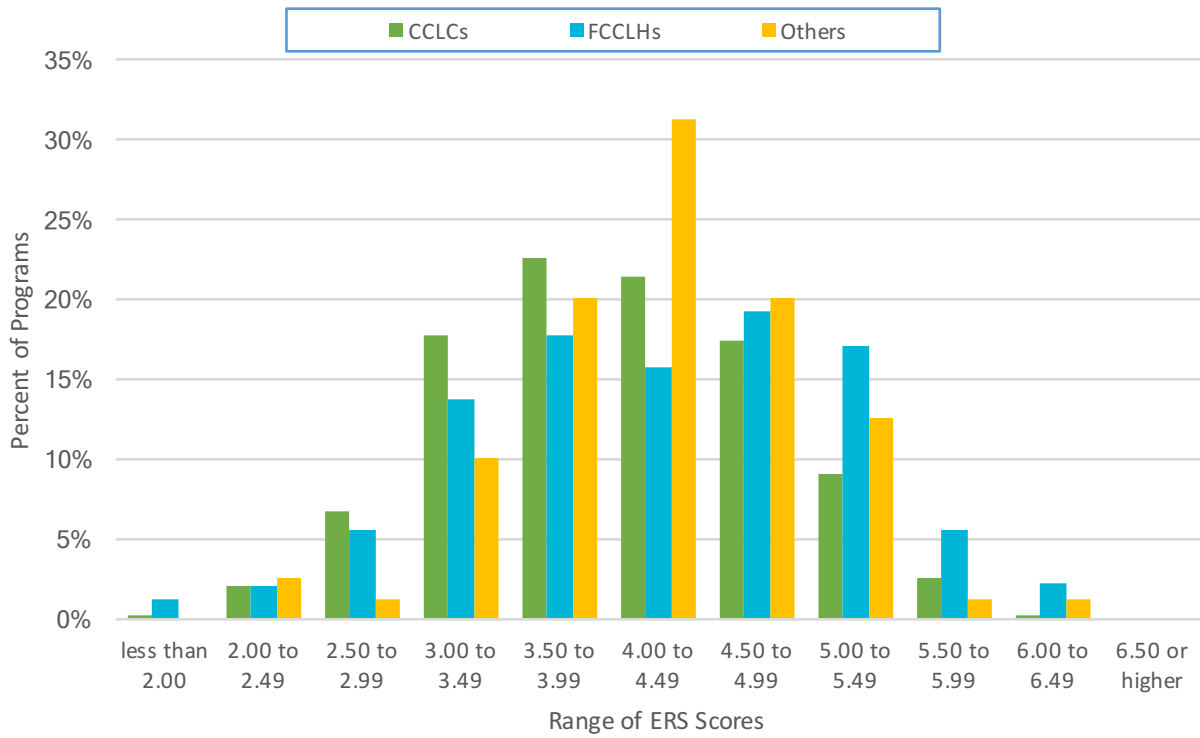
Figure C2 shows the distributions of the ERS scores for the three program types. Whereas the average difference on ERS between CCLCs and FCCLHs was only 0.3, we see from this figure that a higher proportion of FCCLHs than CCLCs attained 5.0 or more on the ERS, explaining why more FCCLHs earned a 3-star rating.

^a The ECERS-3 is the latest version of the ECERS. DECAL transitioned to using it early in 2017.

^b See <http://ersi.info/> for more information about the ERS.

Figure C2. Distribution of average ERS scores for each program type

A higher proportion of FCCLHs and Others than CCLCs had an ERS score above a 5.0.



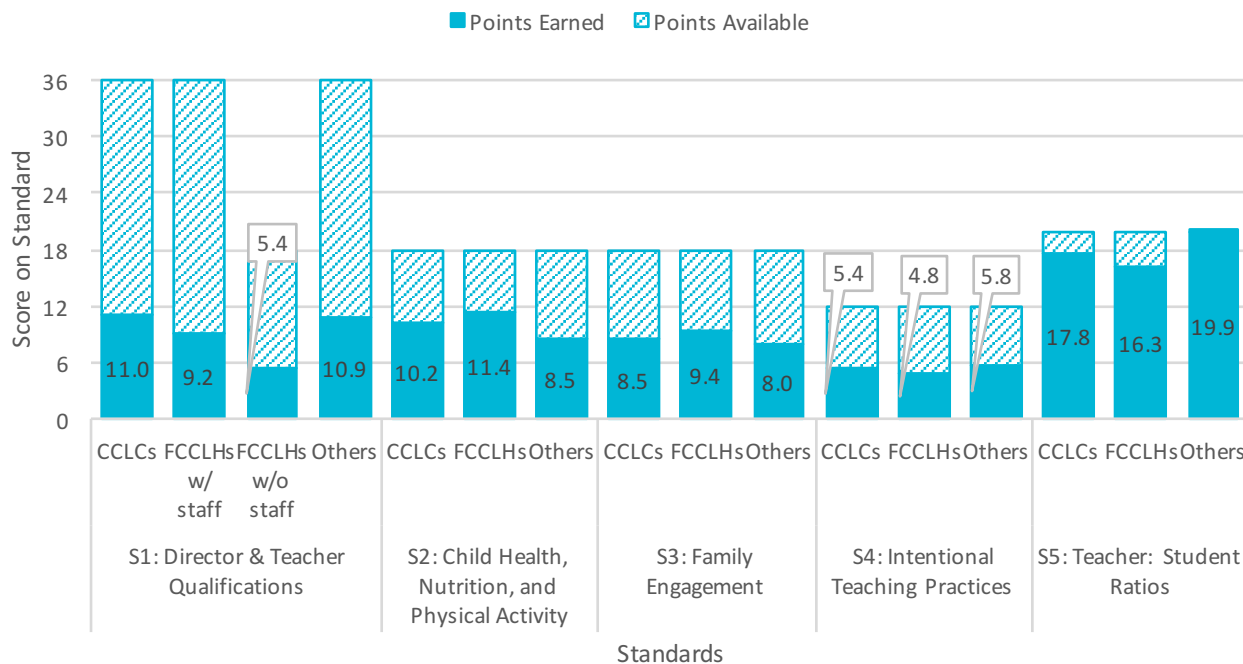
Source: DECAL's administrative data as of May 31st, 2017

Appendix D: What is the distribution of scores for each Quality Rated standard in the portfolio?

Figure D1 shows the average scores each type of program earned for each standard. The striped areas indicate the maximum score available.^a Programs typically earned most of the available *Teacher:Student Ratio* points (Standard 5), but programs earned a lower proportion of possible *Director and Teacher Qualifications* (Standard 1) points. There were many significant differences between program types. Some of the most interesting were that FCCLHs tended to earn more points on *Family Engagement* and CCLCs and Others earned more points on *Intentional Teaching Practices* than FCCLHs.

Figure D1. Average scores earned on each standard for each program type

Programs typically earned a high number of points on the *Teacher:Student Ratio* standard and fewer points on the *Intentional Teaching Practices* standard.



Source: DECAL's administrative data as of May 31st, 2017

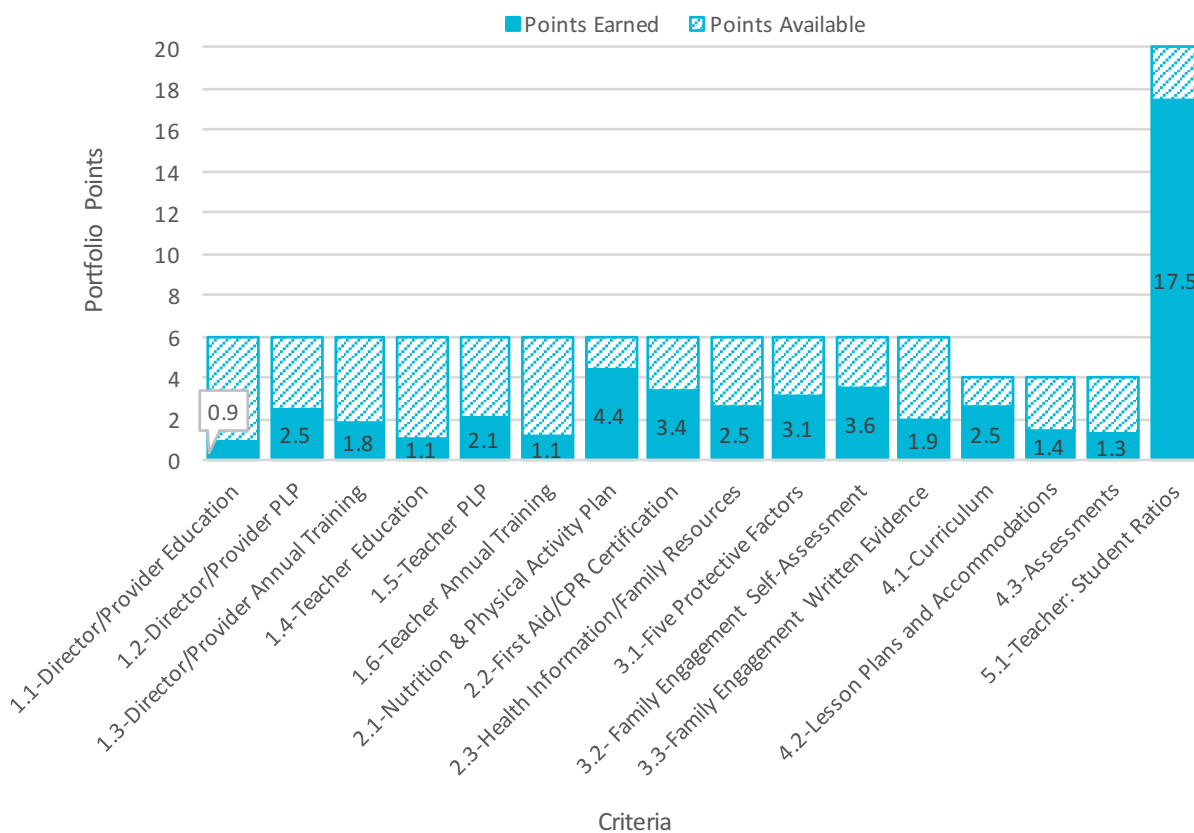
^aNote that this standard combines director and teacher qualifications and, for most programs, 36 points are available. However, only 18 points are available for FCCLH programs without staff because those programs do not have teachers. The rubric used to convert portfolio points to star levels is also different for FCCLH programs without staff, to account for this difference in maximum available points.

Appendix E: What is the distribution of points earned in the individual criteria that make up the Quality Rated standards in the portfolio?

Figure E1 shows the average number of points programs earned on each of the criteria that make up the standards, as well as the number of available points, for all programs. Programs generally earn a high proportion of the available points on *Nutrition and Physical Activity, Curriculum, and Family Engagement Self-Assessment*. The criteria that appear to be most difficult for programs were *Director/Provider Education, Teacher Education, and Teacher Annual Training*.

Figure E1. Average portfolio points for each criteria

Programs attained a high proportion of points for Nutrition & Physical Activity Plan, Curriculum, and Family Engagement Self-Assessment; Education and Training points were more challenging for programs to earn.

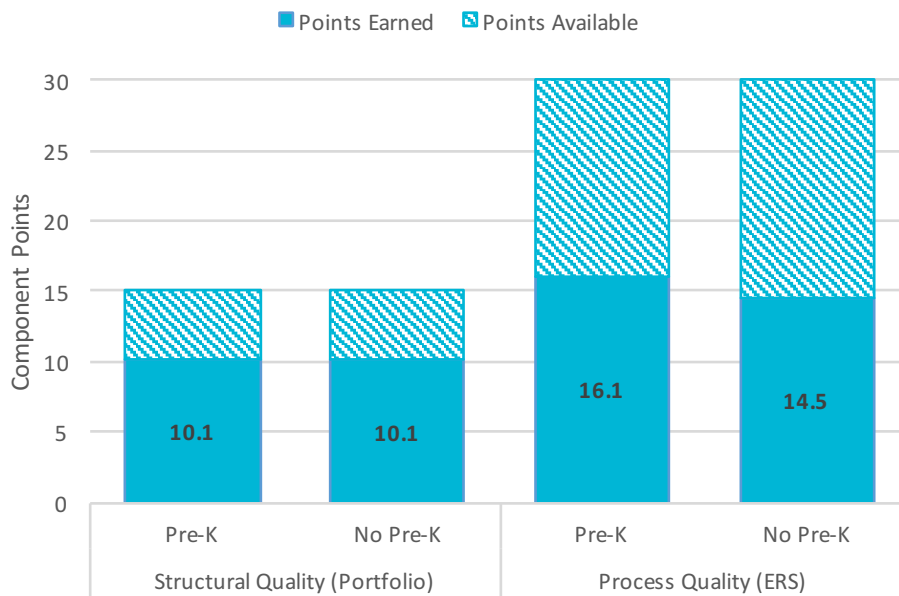


Source: DECAL's administrative data as of May 31st, 2017

Appendix F: How do Child Care Learning Centers with Georgia's Pre-K earn points in Quality Rated?

There were no differences between CCLCs that did and did not include Georgia's Pre-K in terms of Structural Points earned, but CCLCs that included Georgia's Pre-K earned significantly more Process Quality points (see Figure F1).

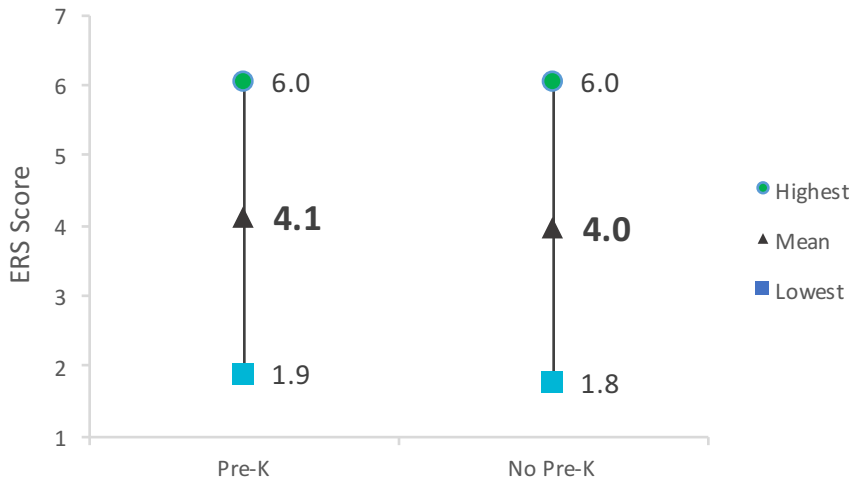
Figure F1. Average Component Points for programs with and without Georgia's Pre-K (CCLCs only)
CCLCs with Georgia's Pre-K earned more Process Quality points.



Source: DECAL's administrative data as of May 31st, 2017

CCLCs with Georgia's Pre-K had significantly higher average ERS scores than CCLCs without Pre-K (see Figure F2). Although the difference between the average ERS scores is only 0.1, 55 percent of CCLCs with Georgia's Pre-K had an average ERS score of 4.0 or higher, compared to only 46 percent of CCLCs without Georgia's Pre-K. This explains how CCLCs with Georgia's Pre-K earn more Process Quality points on average despite only a small average difference on ERS between the two groups.

Figure F2. Average ERS scores for programs with and without Georgia's Pre-K (CCLCs Only)
Programs with Georgia's Pre-K had higher ERS scores.

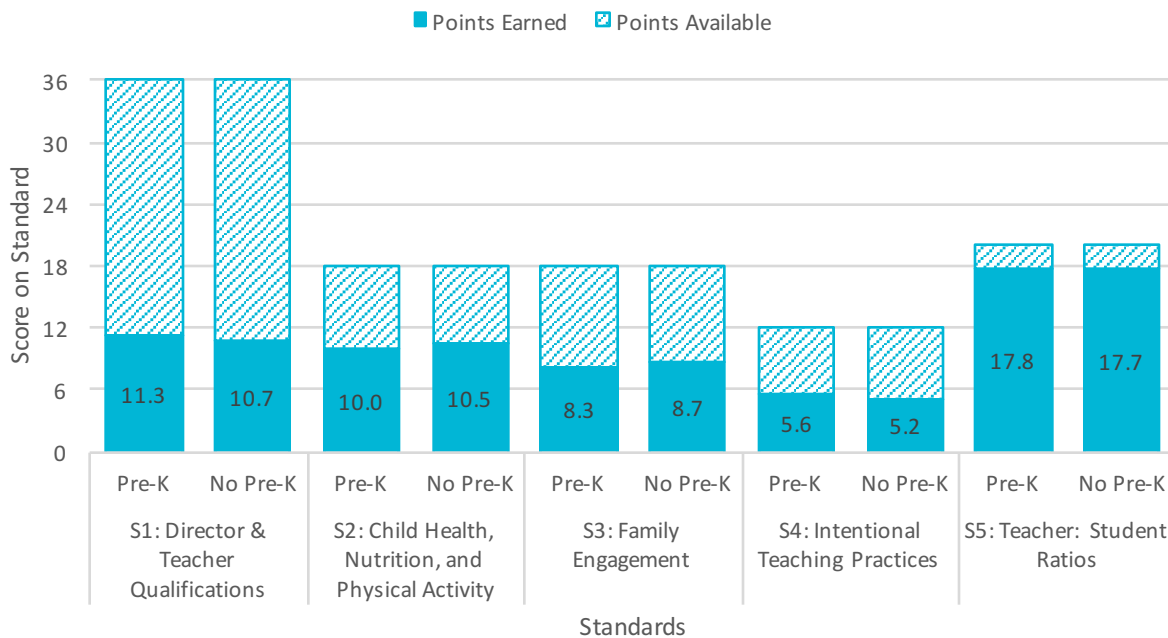


Source: DECAL's administrative data as of May 31st, 2017

There were no significant differences between CCLCs that do and do not include Georgia's Pre-K in any of the individual standards that make up the portfolio (see Figure F3).

Figure F3. Average scores earned on each standard by programs with and without Georgia's Pre-K (CCLCs only)

Scores on standards were similar for CCLCs with and without Georgia's Pre-K.



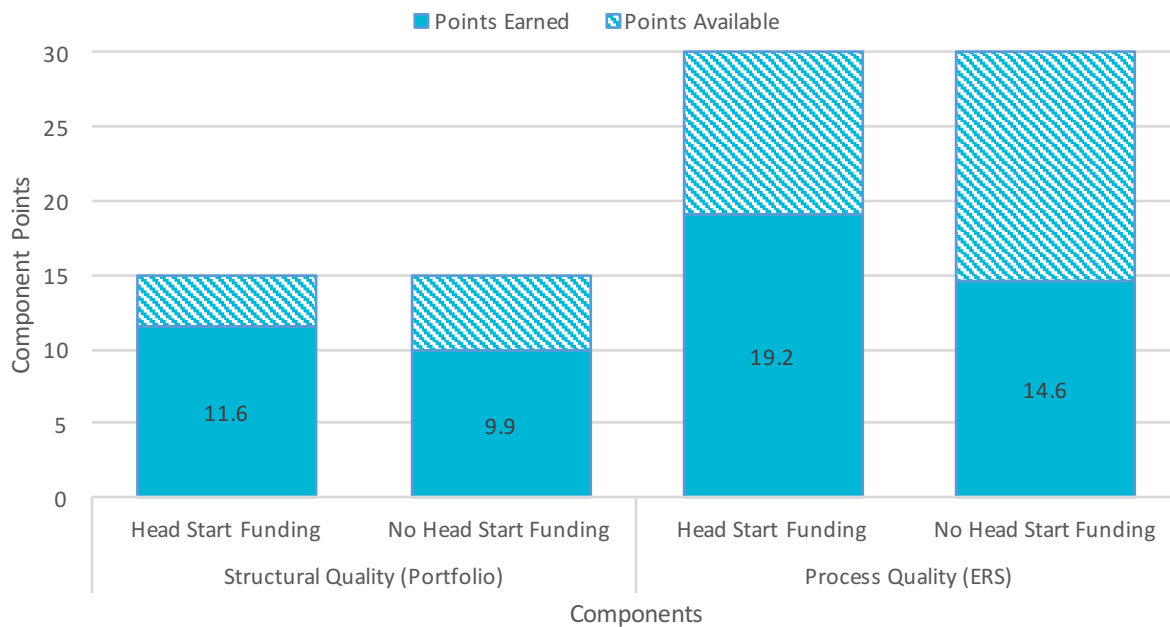
Source: DECAL's administrative data as of May 31st, 2017

Appendix G: How do Child Care Learning Centers and Others with Head Start funding earn points in Quality Rated?

CCLCs and Others with Head Start funding earned significantly more Structural and Process Quality points than those without Head Start funding (see in Figure G1).

Figure G1. Average Component Points earned by programs with and without Head Start funding (CCLCs and Others only)

CCLCs and Others with Head Start funding earned more Structural and Process Quality points than those without Head Start funding.

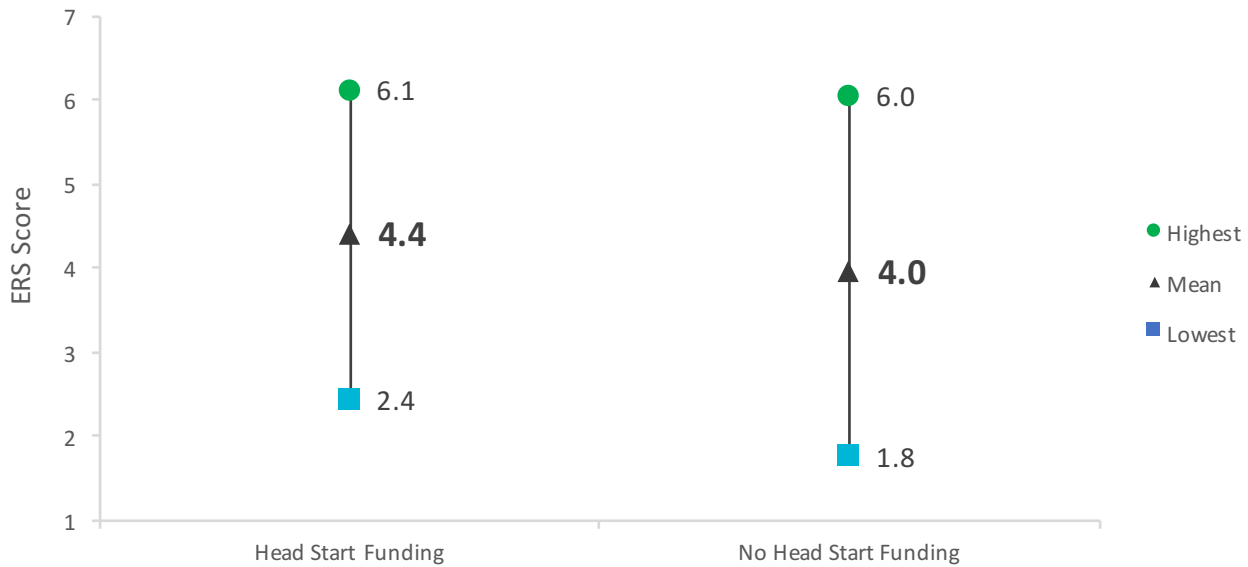


Source: DECAL's administrative data as of May 31st, 2017

CCLCs and Others with Head Start funding had average ERS scores that were significantly higher than CCLCs and Others without Head Start funding (see Figure G2).

Figure G2. Average ERS scores for programs with and without Head Start funding (CCLCs and Others only)

CCLCs and Others with Head Start scored higher on the ERS than those without Head Start funding.

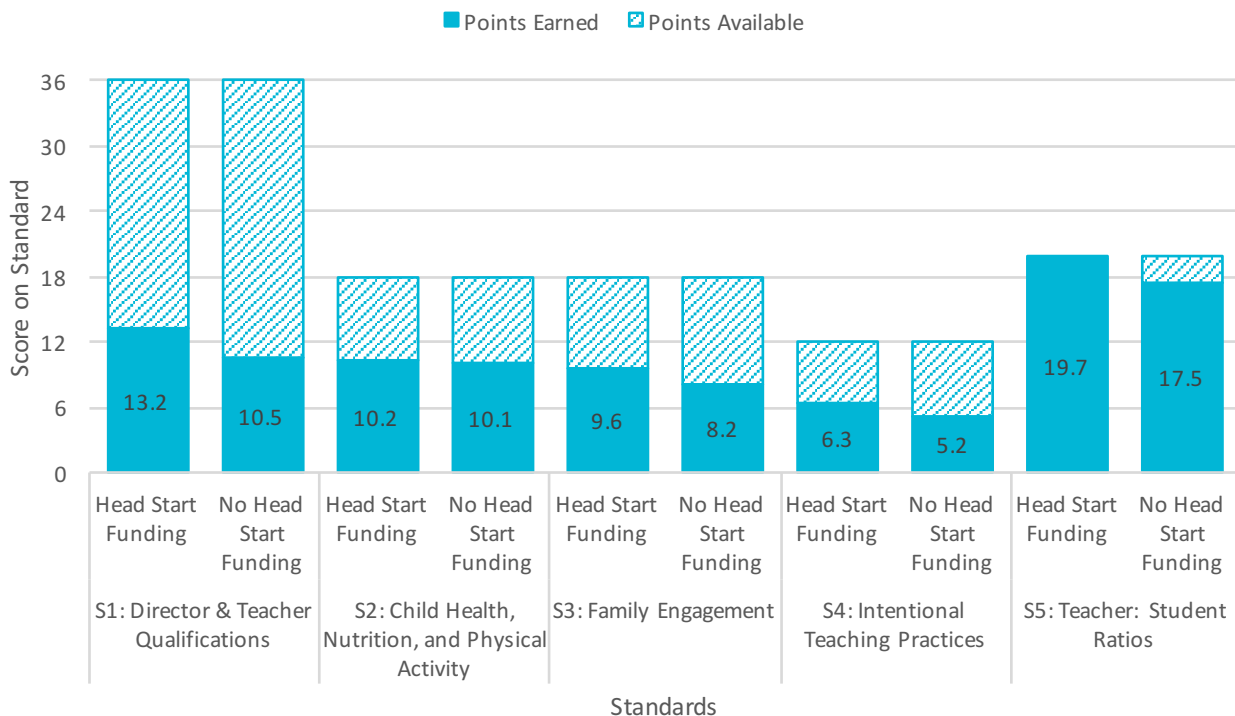


Source: DECAL's administrative data as of May 31st, 2017

Figure G3 shows the average points earned on each standard by CCLCs and Others with and without Head Start funding. CCLCs and Others with Head Start funding scored significantly higher on all standards than those that did not receive Head Start funding, except for the *Child Health, Nutrition, and Physical Activity* standard.

Figure G3. Average scores earned on each standard by programs with and without Head Start funding (CCLCs and Others only)

CCLCs and Others with Head Start funding scored higher than those without on four of five standards.



Source: DECAL's administrative data as of May 31st, 2017

Appendix H: How do the various Quality Rated criteria, standards, and ERS relate to one another?

Quality Rated’s star rating includes a variety of criteria and standards, in addition to the ERS. It is important to understand whether and how these quality dimensions relate to each other and whether they measure a single underlying construct. The odds that a star rating will be related to a more narrowly-defined measure of quality, such as the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) or to children’s development, increases if the star rating measures a single construct of quality (Burchinal, Tarullo, Zaslow, 2016). The following section addresses this question.

Are the criteria within each standard generally measuring a single underlying quality construct?

Table H1 provides Cronbach’s alpha for the criteria that make up each of the standards within the Quality Rated system. Cronbach’s alpha tells us the extent to which the criteria within each standard are measuring a single construct. Alphas can range from 0 (indicating no association among the items) to 1.00 (indicating that all items are measuring exactly the same thing). When creating a measurement scale, an alpha of .70 is generally considered acceptable (Nunnally, 1978), but those guidelines are based on an assumption that the goal is for the scale to measure just one construct. In the case of Quality Rated, DECAL may have intended to combine multiple constructs. Thus, the alpha provides some information about how criteria are functioning from a statistical point of view, but they do not necessarily indicate the extent to which the scale is functioning as DECAL intended.

As seen in Table H1, all alphas are below .70. This suggests that the standards are measuring multiple constructs. Put another way, the *Intentional Teaching Practices* standard includes three different criteria: (1) *Curriculum*, (2) *Activity/Lessons Plans and Accommodations for Every Child*, and (3) *Assessments*. These low alphas show us that a program with a strong *Curriculum* may or may not be strong in *Assessment*; therefore, combining them into a single scale may mask important distinctions. (Note that the final Quality Rated Standard—*Teacher:Student Ratios*—includes only one criteria, so Cronbach’s alpha is not meaningful and could not be calculated.)

Table H1. Cronbach’s alpha for the criteria in each standard

The criteria that make up the standards do not generally measure a single construct.

Standards	Cronbach’s alpha (all programs)
S1A: Director Qualifications	0.48
S1B: Teacher Qualifications	0.56
S2: Child Nutrition and Physical Activity	0.60
S3: Family Partnership	0.68
S4: Intentional Teaching Practices	0.57

Source: DECAL’s administrative data as of May 31st, 2017

Appendix I: Where are 0-star programs typically losing points?

Programs that apply to be rated but do not earn a star are called *0-star* for the purpose of this report. Table I1 shows the average scores of the 0-star programs ($n = 124$) for each standard, as well as their average Structural Quality and Process Quality points. As a comparison, it shows the average scores for 1-star programs ($n = 517$).

Programs with zero stars earned significantly fewer points than 1-star programs on the *Teacher Qualifications*, *Intentional Teaching Practices*, *Teacher:Student Ratios* standards, as well as Structural Quality and Process Quality points; there were no significant differences between 0-star programs and 1-star programs on the other standards. However, ERS scores are clearly the main barrier for 0-star programs. Whereas 0-star programs earned roughly 89 percent as many Structural Quality points as 1-star programs (7.9 vs. 8.9), 0-star programs earned only 1 percent as many Process Quality points (0.1 vs. 9.9). The average Process Quality points of 0.1 is driven by an average ERS score of 2.6 for 0-star programs, which is very low.

Table I1. Scores on standards and components for 0- and 1-star programs

Process Quality points prevent 0-star programs from attaining a star.

Standards and components	0-star program mean	1-star program mean
S1A: Director Qualifications	3.5	3.7
S1B: Teacher Qualifications	2.9	3.7
S2: Nutrition & Activity	9.5	9.4
S3: Family Partnership	6.7	7.7
S4: Intentional Teaching Practices	3.9	4.6
S5: Teacher:Student Ratios	16.2	17.2
Structural Quality Points	7.9	8.9
Process Quality Points	0.1	9.9

Source: DECAL's administrative data as of May 31st, 2017

