Step Up to Quality Validation Study

THE IMPACT OF STEP UP TO QUALITY ON CHILD OUTCOMES, JANUARY 2020

Sheri Chaney Jones, M.A.
Alyssa Pettey, M.S.
Nicholas J. Rockwood, Ph.D.

Measurement Resources Company
MEASUREMENTRESOURCESCO.COM | 614-947-8899

Contents

Executive Summary

Step Up To Quality (SUTQ), Ohio's Quality Rating and Improvement System (QRIS), was implemented statewide in 2006 to measure and promote the quality of all early learning and development programs, including programs in child care facilities, school districts, and family child care homes licensed by the Ohio Department of Job and Family Services (ODJFS) or the Ohio Department of Education (ODE). The most current program standards of SUTQ reflect best practices in early learning environments by encouraging programs to engage in intentional administrative and instructional staff support, screenings, assessments and referral processes, as well as align curriculum to age appropriate Ohio Early Learning and Development Standards.

ODJFS and ODE are committed to a process of evaluation, reevaluation, and refinement in which evidence is continuously gathered to understand whether the inferences drawn from a measure, in this case the SUTQ rating, support its intended purpose. In 2019, a competitive selection process resulted in the engagement of Measurement Resources Company to conduct research assessing the impact of differential levels of program quality, as measured by SUTQ, on Kindergarten Readiness Assessments (KRA) and third grade English Language Arts (ELA) results. Test scores (KRA and third grade ELA) and data of program characteristics (e.g. percent of children enrolled by race, geographic region), child characteristics (race, poverty level, gender, age, and household size), and quality rating scores of the early learning programs the children attended were gathered for all children who attended publicly-funded early learning programs from 2012 through 2019. Analyses revealed that, after accounting for program and child characteristics, early learning program quality, as measured by SUTQ ratings, is related to positive child outcomes. Key findings from these analyses are summarized below.

Key Findings

1. Children who participate in rated programs, on average, score higher on the overall KRA and all KRA subscales compared to their peers who participate in nonrated programs. Children who attended rated programs, on average, had statistically significantly[1] higher KRA scores than their peers in nonrated programs, ranging from 0.17

By the Numbers

Children in 5-Star programs score an average of 1.91 points higher in kindergarten Language and Literacy than children in 1-Star programs

Children in 4-Star programs score an average of 3.36 points higher on third grade ELA than children in 1-Star programs

Children who spend two years in a PFCC rated program score an average of 0.45 points higher on kindergarten readiness than children who spend two years in a PFCC nonrated program

Children in higher-rated programs score an average of 0.81 points higher on Language and Literacy than children in lower-rated programs

---

[1] Statistical significance indicates there is enough evidence to conclude that there are differences in the outcome (i.e. KRA and third grade ELA scores) between the groups being compared (e.g. rated versus nonrated and star ratings). Statistical significance does not indicate the size of the effects or the meaningfulness of such differences. The size of the differences is reported in terms of the number of points groups are expected to differ on the KRA and third grade ELA.

to 0.29 points higher (on a scale ranging from 202 to 298 points) for each year spent in a rated program. These effects are typically considered small.[2]

2. **Children in 5-Star programs scored higher on the KRA Language and Literacy.** Children in 5-Star programs had statistically significantly higher scores than their peers in programs with ratings from 1- to 4-Stars, scoring, on average, 0.91 to 1.91 points higher than their peers; these effects are typically considered small.[3]

3. **Children in 3- and 4-Star programs scored higher on third grade ELA than their peers in 1-Star programs.** Children's scores in 3- and 4-Star programs were statistically significantly higher than children in 1-Star programs scoring, on average, 2.62 to 3.36 points higher (on a scale ranging from 545 to 863 points) than their peers on third grade ELA; these effects are typically considered small.[4]

4. **Children in higher-rated programs consistently scored higher on the overall KRA, KRA Language and Literacy, KRA Mathematics, and third grade ELA when compared to their peers in lower-rated programs.** Program quality was condensed into two categories, lower-rated (1- and 2-Star) and higher-rated (3-, 4-, and 5-Star) programs. Based on these categorizations, children in higher-rated programs scored statistically significantly higher than their peers in lower-rated programs. These effects are typically considered small.[5]

5. **Children's growth between kindergarten readiness and third grade ELA does not vary based on quality rating of the program.** The relationship between KRA and third grade ELA scores based on quality rating were assessed. Overall, KRA scores significantly and positively predicted scores on third grade ELA (higher KRA scores were related to higher third grade ELA scores), however children with the same KRA scores were not estimated to differ in their third grade ELA scores due to the ratings of the early learning programs they attended through third grade.

Overall, this study has shown that early learning program quality, as measured by SUTQ ratings, is related to positive developmental outcomes in children at kindergarten and third grade when examining performance on statewide assessments. There is a consistent, positive relationship between participating in rated programs, regardless of star rating, and kindergarten readiness. Children who participate in rated programs score higher on Ohio's Kindergarten Readiness Assessment than their peers who attend programs that are not rated. Further, children who attended higher-rated (3- to 5-Star) programs score higher on kindergarten readiness and third grade ELA when compared to their peers in lower-rated programs (1- and 2-Star). However, when assessing the extent to which increasingly higher quality (i.e. moving from 1- through 5-Stars) leads to increasingly higher kindergarten readiness and third grade ELA scores, findings are mixed and show that scores do not statistically significantly increase in line with quality rating increases. For example, overall KRA, KRA Language and Literacy, and KRA Mathematics scores increased, on average, as rating level increased, however, many of these differences were not statistically significant. These findings indicate that early learning program quality, as measured by SUTQ ratings, is related to positive child developmental outcomes; however categorizing quality into five levels does not lead to consistent positive trends in child outcomes.

---

[2] Effect sizes for these relationships ranged from R^2 = .0002 to .0006. Effect sizes of .02 or below are considered small according to standards commonly used in the social sciences: Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155-159

[3] The standardized mean difference closely resembles that of Cohen's d (a commonly understood effect size) and is interpreted as such. These effects are classified as small, ranging from .08 to .15, according to standards commonly used in the social sciences (i.e., small = .20, medium = .50, large = .80): Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155-159.

[4] Standardized mean difference for 4-Star to 1-Star = 0.09; Standardized mean difference for 3-Star to 1-Star = 0.07

[5] Standardized mean differences for higher-rated programs compared to lower-rated programs ranged from .04 to .07

Introduction

Step Up To Quality (SUTQ), Ohio's Quality Rating and Improvement System (QRIS), was implemented in child care settings voluntarily in 2006 to measure and promote the quality of early learning and development programs, including programs in child care facilities, and Type A family child care homes licensed by the Ohio Department of Job and Family Services (ODJFS). At that time, SUTQ was a 3-star QRIS system and ODE licensed programs did not participate in the system. In 2011, Ohio was awarded the Early Learning Challenge Grant and Ohio created a plan to expand SUTQ to a 5-star QRIS that included both ODJFS and ODE licensed programs. The current program standards of SUTQ, which were enacted in 2014 as part of Ohio state law for all types of publicly funded early learning programs (i.e. early childhood education programs, preschool special education, publicly funded child care, and publicly funded family child care programs), reflect best practices in early learning environments by encouraging programs to engage in intentional administrative and instructional staff support, screenings, assessments and referral processes, as well as align curriculum to age appropriate Ohio Early Learning and Development Standards. To understand the impact of quality early learning programs on kindergarten readiness and third grade test scores in the state of Ohio, Measurement Resources Company (MRC), in partnership with the Ohio Department of Job and Family Services (ODJFS) and the Ohio Department of Education (ODE), conducted research assessing the impact of differential levels of program quality, as measured by SUTQ, on kindergarten readiness and third grade English Language Arts (ELA) assessments. This report describes data analyses conducted to answer the extent to which program quality, as measured by SUTQ, has a positive impact on child outcomes. The following sections of this report provide an overview of the SUTQ rating system, previous research on quality rating and improvement systems, and analyses conducted to determine the extent to which SUTQ impacts kindergarten readiness and third grade ELA test scores.

Step Up To Quality: Ohio's Quality Rating and Improvement System

Step Up To Quality (SUTQ) is Ohio's tiered quality rating and improvement system which uses a 1- to 5-Star rating system to assess the quality of ODJFS and ODE early learning and development programs. SUTQ uses a hybrid rating structure in which a building-block and points structure are combined. The building-block approach is applied for star ratings 1 through 3 in that to be promoted to the next rating level, all standards in the lower level must be met. Beyond 3 Stars, programs can earn points for meeting additional standards. Programs who participate in SUTQ are rated on four domains, each with subdomains, including learning and development, administrative and leadership practices, staff qualifications and professional development, and family and community partnerships. Moreover, there are two optional additional point areas: staff/child ratios and program accreditation by an approved body, in which programs receive additional points towards a higher rating if the standards are met (Table 1).[6]

Table 1. SUTQ Domains, Subdomains, and Additional Points

Domain 1: Learning and Development
Curriculum and planning
Child screening and assessment
Staff/child interactions and learning environments
Domain 2: Administrative and Leadership Practices
Staff supports and wage structures
Program administration
Staff management
Domain 3: Staff Qualifications and Professional Development
Staff education
Professional development
Domain 4: Family and Community Partnerships
Transition planning for transitions into and out of the program
Communication and engagement with families and community members
Additional Points
Staff/Child ratios
Program accreditation

The 1- to 5-Star ratings are intended to reflect differential levels of quality for each of the domains, with 1-Star representing the lowest quality and 5-Star representing the highest. Each level of quality has specific standards that must be achieved to receive that rating. Additional points are awarded for programs demonstrating lower staff/child ratios than required for program licensure, having more highly educated staff and/or leadership, staff participation in additional professional development, thoroughly defining how curriculum and planning considers individual children's needs, providing children a variety of daily experiences, and ongoing monitoring of child screening and assessments, including aligning screening and assessment scores with instruction. If enough points are earned, a program can earn a 4- or 5-Star rating. More information on the standards at each SUTQ rating level can be found here.

As of 2014, all ODJFS licensed child care centers, ODE licensed preschool programs, and family child care home providers (Type A and Type B) were eligible to participate in SUTQ (ODJFS licensed child care centers and Type A home providers began participating in SUTQ in late 2006). Early childhood education programs were required to be 3-star rated or higher in SUTQ in 2016. In 2018, preschool special education classrooms were required to be star rated. By July 2020, all of the publicly funded child care (PFCC) providers will be required to participate in SUTQ as a condition of funding.[7] Further, by July 2025 all PFCC providers, excluding those small family child care providers (Type B) who are a PFCC provider, will need to be rated at a 3-Star or higher to receive funding.[8]

The state of Ohio is one of 14 states considered as early adopters of a statewide QRIS with the early version of Ohio's QRIS beginning in child care settings voluntarily in late 2006.[9] The current five-star SUTQ system was adopted for all early learning and development programs in 2014. The maturity of the

---

[7] Ohio Revised Code (ORC), section 5104.31
[8] ORC 5104.29
[9] https://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG795.pdf

current SUTQ system, having been in operation for five years, provides robust data that will inform the extent to which quality in early learning environments impacts developmental outcomes in children. The following section discusses recent research on the impact of QRIS on child development in Ohio and other states.

Previous Research on Quality Rating and Improvement Systems and Impacts on Child Development

Research on Ohio's QRIS

In 2016, ODJFS commissioned a study on SUTQ to understand whether the rating and improvement system was achieving its intended outcomes.[10] One focus of this study was the extent to which program quality, as measured by SUTQ rating, impacted Kindergarten Readiness Assessment scores (KRA). The researchers found that when star ratings were collapsed into two categories (lower ratings: 1-and 2-Star compared to higher ratings: 3-, 4-, and 5-Star), children in programs with higher ratings had statistically significantly higher KRA scores than children in the lower-rated programs, though the sizes of the effects found are typically considered small.[11] When KRA scores were compared across four levels of quality (4- and 5-Star programs were combined due to a limited sample size), the only significant difference found was in language and literacy with children who attended 4- and 5-Star programs having higher scores than those who attended 1-Star programs. These findings suggest that, overall, children who attended higher-rated programs in preschool outperformed their peers who did not attend higher-rated programs on Ohio's KRA, yet performance did not always consistently increase as program quality increased.

Previous research on other states' QRIS

Understanding the impact of quality on child outcomes is critical to validating states' QRIS. If QRIS standards are successful at promoting the development of children served, then higher QRIS ratings should lead to positive child developmental outcomes. To understand the relationship between QRIS and child development, several states or sub-state regions have commissioned studies on the impact of QRIS on child outcomes including Colorado,[12] Missouri,[13] Indiana,[14] Minnesota,[15] Delaware,[16] and California,[17] to name a few. Across these studies, the findings are mixed with the majority indicating that there is little evidence to suggest that higher quality predicts better child developmental outcomes. For example, in California, analyses found children's mathematics and early literacy scores were comparable across the four tiers of quality and in another measure of early literacy, children in lower-rated programs actually outperformed children in higher-rated programs.[18] Similarly, in a study conducted in Virginia,[19] children in lower-rated programs had statistically significantly higher literacy scores compared to children in higher-rated programs. Despite children in higher-rated programs having lower scores in

---

[10] http://earlychildhoodohio.org/sutq/pdf/SUTQValidationStudy2017.pdf

[11] The largest effect size obtained was differences in overall kindergarten readiness scores for 2015-2016 with an effect size (Cohen's d = .28) generally interpreted as a "small" effect.

[12] https://www.rand.org/pubs/monographs/MG650.html

[13] https://www.marc.org/Community/Early-Learning/pdf/QRSfindings.aspx

[14] https://www.in.gov/fssa/files/PTQFinalReportRev11012.pdf

[15] https://www.childtrends.org/wp-content/uploads/2016/03/Parent-Aware-Validation-and-Letter-3-2016.pdf

[16] https://www.mathematica.org/our-publications-and-findings/publications/are-ratings-from-tiered-quality-rating-and-improvement-systems-valid-measures-of-program-quality

[17] https://www.air.org/sites/default/files/downloads/report/RTT-ELC%20QRIS%20Cumulative%20Technical%20Report%20-%20FINAL.pdf

[18] https://www.air.org/sites/default/files/downloads/report/RTT-ELC%20QRIS%20Cumulative%20Technical%20Report%20-%20FINAL.pdf

[19] https://www.acf.hhs.gov/opre/resource/validation-of-quality-rating-and-improvement-systems-qris-examples-from

literacy, the authors of the study did find that children in higher-rated programs showed greater developmental gains than children in lower-rated programs.

Overall, across the studies examined, higher ratings in the various QRIS tended to result in higher quality as measured by observations; however, despite the consistent, positive relationship between higher quality as measured by QRIS and higher quality as measured by other measures (e.g. quality of environment, instruction, staff-child interactions), these same studies provided little evidence to conclude that the differences in quality produced meaningful differences in child developmental outcomes.

Across the studies, authors provided several possible explanations for why higher quality, as measured by QRISs, did not lead to consistent, positive outcomes for children, ranging from study design limitations to methodological challenges of the QRISs themselves. Study design limitations included low statistical power as a result of small sample sizes and sample selection biases (i.e. programs self-selected to participate in some studies). Possible QRIS methodological explanations for the lack of consistent associations between quality and child outcomes include the broad way in which quality is measured under many of the QRISs. Authors explained that because quality is often measured in broad categories made up of several specific indicators, there are likely measurement issues impacting examinations of the relationship between quality and child outcomes. For example, in Ohio, the Administrative and Leadership Practices domain is comprised of three specific indicators (staff supports, program administration, and staff management). Because these three indicators likely vary in their importance to child outcomes, collapsing ratings across the three into one indicator rather than looking at the impact of each individually, attenuates the relationship between quality and child outcomes. Consequently, the way in which quality is measured may be contributing to methodological limitations which limit our ability to identify how each component of quality impacts child outcomes. Another explanation offered by the authors is that the different levels of quality (e.g. a 1-Star compared to a 2-Star) may not be large enough to produce differences in child outcomes. That is, the different requirements to become a higher-rated program may not be enough to affect changes in child developmental outcomes.

The present study builds upon this previous research by examining the impact of SUTQ on kindergarten readiness and third grade English Language Arts (ELA) test scores as well as child growth between KRA and third grade ELA.

Methodology

The present study included child and early learning program data gathered from ODJFS and ODE from 2012 through 2019. ODJFS provided data for all children who attended PFCC from 2012 to 2019. ODE provided KRA and third-grade ELA testing data for all children (regardless of whether the children attended PFCC) from 2015 to 2019. The data provided were used to answer the following research questions:

- What is the relationship among program rating or nonrated status and outcomes of children who are publicly funded, as measured by kindergarten readiness?
- Do kindergarten readiness assessment test scores vary by quality rating?
- Do third grade ELA test scores vary by quality rating?
- Does a child's progress (i.e. from KRA to third grade ELA) vary by quality rating?

Because the current five-star SUTQ rating system did not take effect until October 2013, only program rating data from this point forward are considered in the analyses. For non-program rating data (e.g. number of months enrolled in early learning programs, percent of children identified as a racial minority enrolled in a program), data from 2012 onward are included. Children's corresponding program ratings were only included for programs attended prior to the date in which they took the KRA and/or third grade ELA in the respective analyses.

Sample

Statewide student identifiers (SSID) were provided from ODJFS for all children who attended a PFCC program from 2012 to 2019, including the associated program the child was enrolled in each month. Only children enrolled in PFCC programs from October 2013 through November 2019 with a corresponding KRA and/or third grade ELA test score were included in the SUTQ analysis to align with the current version of SUTQ. Test data for children who did not attend PFCC (i.e. who had testing data from ODE but did not attend PFCC) were only examined to show mean score differences for students who did and did not attend PFCC (see Figure 1 and Table 4 on page 12); these data were not included in the subsequent SUTQ analysis as they were not enrolled in PFCC. A total of 89,691 children who attended PFCC and had KRA and/or third-grade ELA test scores were included in the sample for the SUTQ analysis, though some analyses were reduced to only children who attended rated programs and only children with third grade ELA scores. Consequently, sample sizes for analyses varied based on the research questions.

Variables and Statistical Methods

Variables

Kindergarten Readiness
Kindergarten readiness as measured by Ohio's Kindergarten Readiness Assessment (KRA) is comprised of an overall score and four subdomains: Language and Literacy, Mathematics, Social Foundations, and Physical Well-Being and Motor Development. Using the overall score, children's scores are classified into one of three categories:[20]

- Demonstrating readiness (scores of 270 to 298): The child demonstrates foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.
- Approaching readiness (scores of 258-269): The child demonstrates some foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.
- Emerging readiness (scores of 202-257): The child demonstrates minimal foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.

Third Grade English Language Arts
Similar to KRA scores, third grade English Language Arts (ELA) scores are classified into one of five categories based on the child's score. The categories include:
- Advanced (scores of 752-863)
- Accelerated (scores of 725-751)
- Proficient (scores of 700-724)
- Basic (scores of 672-699)

---

[20] http://education.ohio.gov/getattachment/Topics/Early-Learning/Kindergarten/Ohios-Kindergarten-Readiness-Assessment/Kindergarten-Readiness-Assessment-for-Teachers/OH_KRA_ISR_2019.pdf.aspx?lang=en-US

- Limited (scores of 545-671)

Table 2 displays a list of the variables included in the statistical models. Appendix A provides a more detailed explanation of the variables.

Table 2. Variable List

Variable	Source
Predictor Variables	
Program star rating while the child was enrolled in the program(s) prior to the KRA (for KRA analyses) and/or third grade ELA test administrations (for ELA analyses)	ODJFS
Program rated or unrated status at the time of enrollment	ODJFS
Outcome Variables	
Third grade ELA scores	ODE
KRA scores (Language and Literacy, Mathematics, Social Foundations, Physical Well-Being and Development, and overall KRA score)	ODE
Covariates/Control Variables	
Program region classification as metropolitan, micropolitan, or rural (based on zip code)	ODJFS & Census Bureau
Months child was enrolled in PFCC in Ohio	ODJFS
Program type (Type A, Type B, Child Care Center, or ODE program)	ODJFS
Child's age	ODJFS
The year the KRA and third grade ELA assessments were taken	ODE
Child's average federal poverty level (FPL) during time enrolled in PFCC	ODJFS
Average FPL among children receiving PFCC enrolled in a program	ODJFS
Child's gender	ODJFS
Child's race	ODJFS
Percent of racial minorities enrolled in a program (children receiving PFCC only)	ODJFS
The school district the child was enrolled in during the time the KRA or third grade ELA was taken	ODE
Mobility between kindergarten and third grade (number of different school districts attended)	ODE

Statistical Methods

Linear mixed models accounting for the nested structure of children within early learning programs and within school districts (i.e. school districts the children were enrolled in at the time of the KRA and third grade ELA administrations) were conducted to answer the research questions. That is, students were nested both within their early learning programs and the school districts in which their KRA and third-grade assessments were taken. Additionally, covariates (i.e., control variables) expected to predict differences in KRA and third grade ELA scores were also included in each model.

Quality was assessed in two ways. First, analyses were conducted to determine whether test scores differed across the five levels of quality. Second, analyses compared 1- and 2-Star programs to 3-, 4-, and 5-Star programs. Since all programs except small family child care Type B programs will be required

to achieve a rating of 3-Star or higher as a condition of funding in 2025, quality was assessed in two categories, 'Under 3' and '3 and Over.' The details of each model are explained below followed by a description of the covariates and their relationships with KRA and third grade ELA scores.

Testing whether KRA and third grade ELA scores differ by quality rating

Linear mixed models were used to test for differences in KRA and third grade ELA scores by program rating. Because children may have attended multiple programs, the program specific variables (i.e., program rating, percent of children identified as a racial minority in the program, average federal poverty level among children enrolled, and metropolitan/micropolitan/rural designation) used to predict the children's KRA and third grade ELA scores were weighted based on when the child attended a specific program and the proportion of time that each child spent in the program prior to the KRA and/or third grade ELA test administration (see, e.g., Fielding & Goldstein, 2006 for more information on the weighting approach used in multilevel multiple membership models).[21] For children included in both the KRA and third grade ELA analyses, separate weighted variables were computed for their time spent in programs before the KRA and third grade ELA test administrations, respectively (e.g. a child would have two weighted variables for the percent of children identified as a racial minority, one for KRA analyses and one for third grade analyses). For example, if, prior to the KRA administration, a child spent half of their time in a program with 22 percent of children identified as a racial minority (at the time they were in the program) and the other half in a program with 28 percent of children identified as a racial minority, the child's weighted covariate for percent of children identified as a racial minority in the KRA analyses would be 25 percent; for third grade ELA analyses, this child's weighted covariate for percent of children identified as a racial minority would be adjusted to include data from programs attended after the KRA administration and before the third grade ELA administration. Thus, the weighted covariates consider the dynamic nature of the program specific variables (including program ratings, which may have changed over time).

Testing whether a child's progress varies by program rating

To determine whether a child's progress (from KRA to third grade) varies by program ratings attended through third grade, a similar model as that described above was used to predict third grade ELA scores with KRA scores as an additional covariate. By controlling for KRA scores, differences in third grade ELA scores between the different program ratings correspond to differences in growth/change from KRA to third grade ELA.

Testing the relationship among program rated and nonrated status and KRA scores

Differences in program ratings (the methods described above) were tested using only the children who had participated in at least one rated program. To test for differences in KRA scores between rated and nonrated programs, all children who had attended a publicly-funded child care program (PFCC; both rated and nonrated) were included in the analysis, with the number of months the child had attended a rated program as an additional predictor. Because the total number of months a child had been in any PFCC program was included as a covariate, the new predictor tests whether children that have been enrolled in PFCC the same amount of time tend to differ in their KRA scores based on the amount of time they spent in a rated program.

---

[21] Fielding, A., & Goldstein, H. (2006). Cross-classified and multiple membership structures in multilevel models: An introduction and review.

The following section summarizes descriptive KRA and third grade ELA data for students throughout the state of Ohio, including those who did and did not attend PFCC.

Descriptive Data

Overall, children in PFCC have slightly lower KRA scores, on average, when compared to children not in PFCC. As shown in Figure 1, all children in Ohio whether in PFCC or not, score on average in the "Approaching Readiness" category. Figure 1 displays the average KRA and domain scores for children from 2014 through 2019 and Table 4 displays the associated sample sizes and standard deviations.

Figure 1. Average KRA Scores by PFCC Status
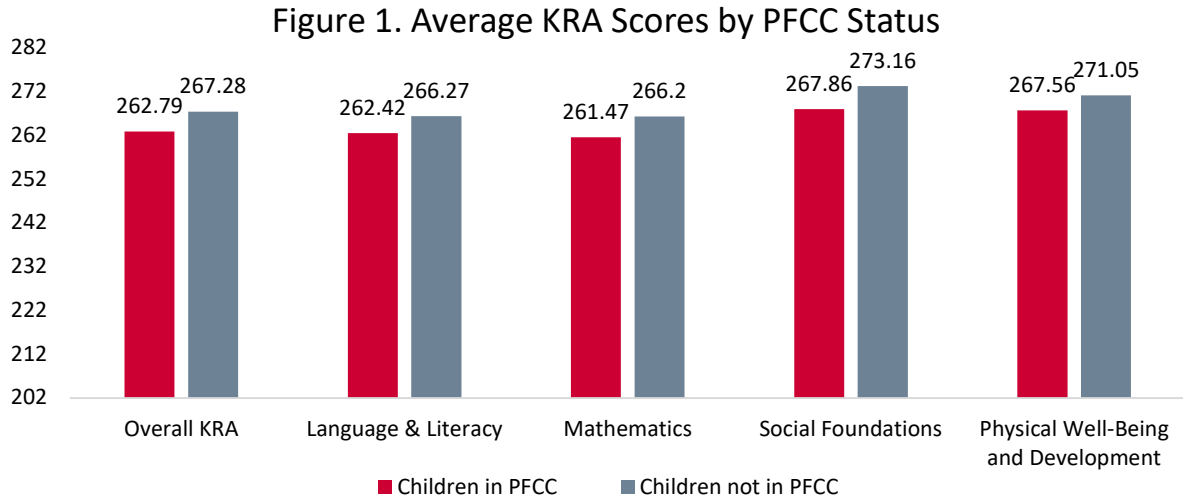
Table 4. Sample size and standard deviations for KRA means in Figure 1.

		Overall KRA	Language & Literacy	Mathematics	Social Foundations	Physical Well-Being & Development
PFCC	**n**	92,766	93,203	93,247	93,016	93,064
	SD	12.00	12.79	12.84	20.55	18.18
Not PFCC	**n**	495,818	498,286	498,755	497,847	497,422
	SD	13.6	14.23	14.69	20.63	18.10

As shown in Figure 2, children who did not attend PFCC have higher third grade ELA scores on average. Further, students, on average score in the "Proficient" range regardless of if they attended PFCC or not.

Figure 2. Average Third Grade ELA Scores by PFCC Status

Figure 2. Average Third Grade ELA Scores by PFCC Status

- Children in PFCC (n = 33,791; SD = 39.06)
- Children not in PFCC (n = 475,038; SD = 47.18)
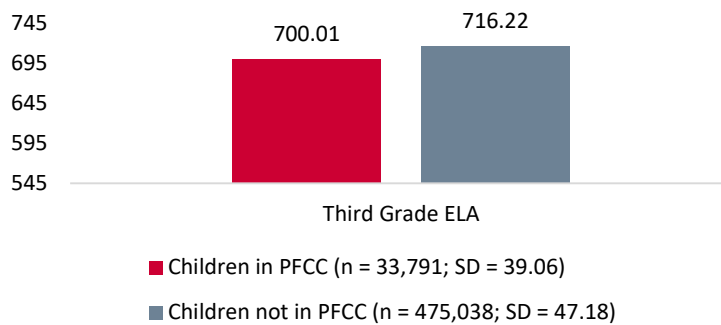
Covariate Relationships

Several covariates were included in the statistical models in order to account for their impact on the outcome variables. Across the models, the covariates had similar relationships with the outcome variables (i.e. KRA and third grade ELA). The general trends are described here; all relationships reported are after accounting for all other variables in the models. For the specific effects of each covariate in each model, please refer to the model statistics found in Appendix B. Table 3 displays the effect sizes for the covariates for the analyses comparing all five levels of program quality. All years of covariate data provided were included in the analyses up until the test administration date. That is, all data available for a child prior to the KRA administration and/or the third grade administration were included in the analyses.

Program type (Type A, Type B, ODE Program, or Child Care Center). Previous research has shown that children who attend child care centers tend to have better academic outcomes compared to children in home-based care.[22] Potential explanations to this effect include more highly educated staff and more structured learning activities in child care centers compared to home-based providers.[23] The results of the current study align to previous research; children who attended Type A and Type B family child care homes consistently scored statistically significantly lower on the KRA and third grade ELA compared to children who attended Child Care Centers, with one exception; children from Type B homes did not score statistically significantly lower on the third grade ELA test than children in Centers. There were not significant differences in the KRA and third grade ELA scores between ODE programs and Child Care Centers.

Gender. Gender has been found to be an important predictor of academic outcomes. Research has shown that girls tend to outperform boys on academic outcomes.[24] In line with previous research, girls in this study outperformed boys in all KRA domains and third grade ELA scoring statistically significantly higher.

---

[22] Bradley RH, Vandell DL. Child care and the well-being of children. Archives of Pediatric and Adolescent Medicine. 2007;161:669–676. doi: 10.1001/archpedi.161.7.669.]

[23] Fuller, B., Kagan, S. L., Loeb, S., & Chang, Y. W. (2004). Child care quality: Centers and home settings that serve poor families. Early Childhood Research Quarterly, 19(4), 505-527.

[24] Duckworth, A., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. Journal of Educational Psychology, 98,198–208

Child's racial minority status. The test performance gap between White and non-White students, particularly Black/African American and Hispanic students, has persisted for decades. Possible explanations for the test score gap and variables not accounted for in these models include cultural differences associated with chronic poverty, disparities in school resources between predominately White schools and predominately minority or Black/African American and Hispanic schools, and differing family dynamics.[25] Children identified as non-White tended to have lower KRA and third grade ELA scores, on average, with the exception of Physical Well-Being and Development.

Percent of children identified as a racial minority enrolled in a program. Because child outcomes are not only influenced by their own characteristics, but also by those of their peers and the environments they are in,[26] the percent of children identified as a racial minority enrolled in a program was also assessed. Similar to the results found for individual child-level racial minority status, programs with higher percentages of children identified as a racial minority tended to have lower KRA and third grade ELA scores, on average. As the percent of children identified as a racial minority among students receiving PFCC enrolled in a program increased, KRA and third grade ELA scores decreased.

Child's average federal poverty level (FPL) during their enrollment in PFCC. The effect of poverty on academic outcomes has long been documented in research; poverty has a strong, adverse impact on students' achievement due to the lack of resources needed for student success.**[27]** In line with previous research, a child's average FPL during their time in PFCC consistently had a significant, positive relationship with KRA and third grade ELA. As the federal poverty level increased (i.e. less poverty), scores increased, on average.

The average federal poverty level among all children receiving PFCC enrolled in each program. As in racial minority status, the analysis not only accounted for a child's individual poverty level, but the poverty level of their peers. Overall, the average federal poverty level among children enrolled in each program consistently had a significant, positive relationship with KRA and third grade ELA. As the average federal poverty level increased (i.e. less poverty), average scores also increased.

Metropolitan, micropolitan, or rural designation based on program zip code. Whether a child lives in a metropolitan, micropolitan or rural area can be an important predictor of academic outcomes as these settings typically differ in terms of their access to resources and cultural norms which may impact child development**.** Research on the effects of location on child outcomes in terms of the urban-rural continuum is rather limited, with studies showing mixed findings. For example, Miller and Votruba-Drzal (2013)[28] found that children in rural areas began kindergarten less advanced than their peers in small urban and suburban areas with the difference largely explained by differing home environments (e.g. rural children had less stimulating environments) and more frequent use of home-based child care rather than center-based. On the other hand, Khattri et al. (1997)**[29]** found that while students from low socioeconomic statuses (SES) in rural areas performed low on achievement measures, their overall

[25] https://www.brookings.edu/articles/the-black-white-test-score-gap-why-it-persists-and-what-can-be-done/

[26] Elias, M. J., & Haynes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. School psychology quarterly, 23(4), 474.

[27] Lacour, M., & Tissington, L. D. (2011). The effects of poverty on academic achievement. Educational Research and Reviews, 6(7), 522-527.

[28] Miller, P., & Votruba-Drzal, E. (2013). Early academic skills and childhood experiences across the urban-rural continuum. Early Childhood Research Quarterly, 28(2), 234-248. doi:10.1016/j.ecresq.2012.12.005

[29] Khattri, N., Riley, K. W., & Kane, M. B. (1997). Students at risk in poor, rural areas: A review of the research.

performance was better than students from comparable SES in urban areas. In the present study, the rural to metropolitan designation had a small impact on how a child performed on the KRA. Children in metropolitan and micropolitan areas tended to score higher than children in rural areas, though the differences tended to be small and only significant for overall KRA scores, Language and Literacy, Mathematics, and Social Foundations. There were not statistically significant differences in third grade ELA scores based on location.

Number of months a child was enrolled in PFCC. The length of time a child spends in child care has been found to predict child academic outcomes, with longer time in care associated with improved academic outcomes.[30] As expected, the number of months a child was enrolled in PFCC statistically significantly and positively predicted KRA scores; such that the more months enrolled, the higher the KRA score. Interestingly, for third grade ELA, there was a negative relationship between months enrolled and third grade scores; longer times enrolled in PFCC were related to reduced scores. This is likely a result of having a restricted dataset given that the PFCC data were limited in time, as data were provided from 2012 to 2019. Thus, student program attendance prior to 2012 is excluded from this analysis.

Child's average household size during their enrollment in PFCC. A child's household size may be a predictor of academic outcomes as family size can impact the amount of social, financial, and academic resources available for each child. Research has found that larger family sizes have a small, negative impact on academic outcomes.[31] In line with these findings, child's household size was negatively related to KRA (all domains except Social Foundations and Physical Well-Being and Development) and third grade ELA scores; as household size increased, scores tended to decrease.

The year the child took the KRA or third grade ELA. To account for year-to-year overall fluctuations in test scores (i.e. the fact that in a given academic school year test scores statewide can be better or worse than the previous year), the year a child took the test was included as a predictor. The year in which a child took the exam was a negative predictor of scores in overall KRA, Language and Literacy, and Mathematics but a positive predictor in Social Foundations. This indicates that, on average, scores in overall KRA, Language and Literacy, and Mathematics have decreased over time, while scores in Social Foundations have improved with time. The year a child took the exam was not a significant predictor of Physical Well-Being and Development or third grade ELA.

---

[30] Lee, K. (2011). Impacts of the duration of Head Start enrollment on children's academic outcomes: moderation effects of family risk factors and earlier outcomes. Journal of Community Psychology, 39(6), 698-716.
[31] Åslund, O., & Grönqvist, H. (2010). Family size and child outcomes: Is there really no trade-off? Labour Economics, 17(1), 130-139. doi:10.1016/j.labeco.2009.05.003

The child's age at the time the KRA or third grade ELA was administered. Child's age at the time the KRA was administered was a strong, positive predictor of KRA scores in that the older a child was, the better he or she performed. This relationship is as expected given the KRA is measuring skills, knowledge, and behaviors children learn as they develop. Age was not a significant predictor of third grade ELA scores.

Mobility between kindergarten and third grade. School mobility has been found to have a negative impact on academic outcomes.[32] In the present study, the number of different school districts a child had been enrolled in between kindergarten and third grade had a statistically significant, negative impact on third grade ELA scores. More school districts attended predicted lower third grade ELA scores, on average.

Table 3. Covariate effects (regression coefficients) on KRA and Third Grade ELA (5-Star Analyses)

	KRA Overall	KRA Language & Literacy	KRA Math	KRA Social Foundations	KRA Physical Well-Being & Development	Third Grade ELA
Program Type A (Reference = Child Care Center)	-1.88***	-1.95***	-1.80**	-2.37**	-2.05**	-10.83***
Program Type B (Reference = Child Care Center)	-1.51***	-1.83***	-1.65***	-1.18**	-1.63***	-2.45
ODE Program (Reference = Child Care Center)	-0.58	.10	-1.30	-.88	-1.41	4.05
Program % Racial Minority Enrolled	-.02***	-.01**	-.02***	-.02**	-.02***	-.13***
Program Average FPL for Children Enrolled	.04***	.04***	.04***	.06***	.05***	.15***
Program Region: Micropolitan (Reference = Rural)	1.25*	1.40*	1.71**	1.38	1.16	1.25
Program Region: Metropolitan (Reference = Rural)	1.46**	1.39*	1.83**	1.86*	1.48	3.05
Child's # of Months Enrolled in PFCC in Ohio	.06***	.07***	.06***	.05***	.06***	-.04*
Child's Minority Status (1 = Minority, 0 = Non-Minority)	-1.02***	-1.21***	-1.72***	-.74**	0.34	-6.13***
Child's Average Household Size while Enrolled in PFCC	-.36***	-.63***	-.40***	-.13	.05	-1.25***
Child's Average FPL while Enrolled in PFCC	.01***	.02***	.01***	.01***	.01***	.05***
Year the KRA or Third Grade ELA was Taken	-.53***	-.89***	-1.04***	.51***	.05	.11
Child's Gender (1 = Males, 0 = Females)	-3.67***	-2.46***	-1.44***	-7.57***	-7.91***	-6.87***
Child's Age when Test was Taken	2.48***	2.49***	2.84***	2.29***	2.29***	.24
Mobility between kindergarten and third grade	--	--	--	--	--	-2.27***

*Note. Effects for lower-higher analyses were comparable in size and direction. Significant codes: * = p < .05, ** = p < .01, *** = p < .001, ns = nonsignificant*

[32] Gruman, D. H., Harachi, T. W., Abbott, R. D., Catalano, R. F., & Fleming, C. B. (2008). Longitudinal effects of student mobility on three dimensions of elementary school engagement. Child development, 79(6), 1833-1852.

Analysis and Results

What is the relationship among program rating or nonrated status and outcomes of children who are funded by publicly funded child care, as measured by kindergarten readiness?

After accounting for child and program characteristics, children who participate in rated programs, on average, score statistically significantly higher on the overall KRA and all KRA subscales compared to their peers who participate in nonrated programs. According to commonly used standards in the social sciences, these are typically considered small effects.[33] For each additional month a child spends in a rated program compared to a nonrated program, they are expected to score 0.014 to 0.024 points higher on the overall KRA and the KRA domains. That is, for two children who attend a program for two years before kindergarten, one child in a rated program and the other in a nonrated program, the child who attended the rated program is predicted to score 0.45 points higher on the overall KRA compared to the child in the nonrated program. This indicates that participating in rated programs, regardless of the quality rating, has a small, positive, and cumulative impact on KRA scores. Table 5 displays the predicted mean score differences for each subscale for every one month, 12 month, and 24 months a child spends in a rated program compared to a nonrated program. Figure 3 displays the adjusted mean KRA scores (i.e. mean scores after removing variability accounted for by the covariates in the models) based on the amount of time a child spends in PFCC. Figure 3 shows the average scores on the KRA and KRA domain scores become slightly higher for each additional 12 and 24 months spent in a PFCC center compared to a nonrated PFCC. The multilevel model statistics are reported in Appendix B (Models 1.1 to 1.5).

Table 5. Predicted mean score increases on KRA for participating in a rated versus nonrated program

KRA Domain	Predicted Mean Score Increase for every 1 Month in a Rated Program	Predicted Mean Score Increase for 12 Months in a Rated Program	Predicted Mean Score Increase for 24 Months in a Rated Program
Overall KRA	0.019	0.227	0.454
Language & Literacy	0.024	0.294	0.588
Mathematics	0.020	0.237	0.474
Social Foundations	0.017	0.202	0.403
Physical Well-Being & Development	0.014	0.172	0.344

---

[33] Number of months in a rated program explained R^2 = .0002 to .0006 additional variance in KRA scores above and beyond the other predictors in the models. Effect sizes of .02 or below are considered small effects according to standards commonly used in the social sciences: Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155-159.

Figure 3. KRA Overall and Domain Scores by Time Spent in a Rated Program Versus a Nonrated Program , Adjusted Means

Legend: 0 Months, 12 Months, 24 Months, 36 Months

Overall: 265.2, 265.4, 265.6, 265.8
Language and Literacy: 264.3, 264.6, 264.9, 265.2
Social Foundations: 270.9, 271.1, 271.3, 271.5
Mathematics: 263.6, 263.8, 264.0, 264.3
Physical Well-Being and Motor Development: 271.0, 271.1, 271.3, 271.5
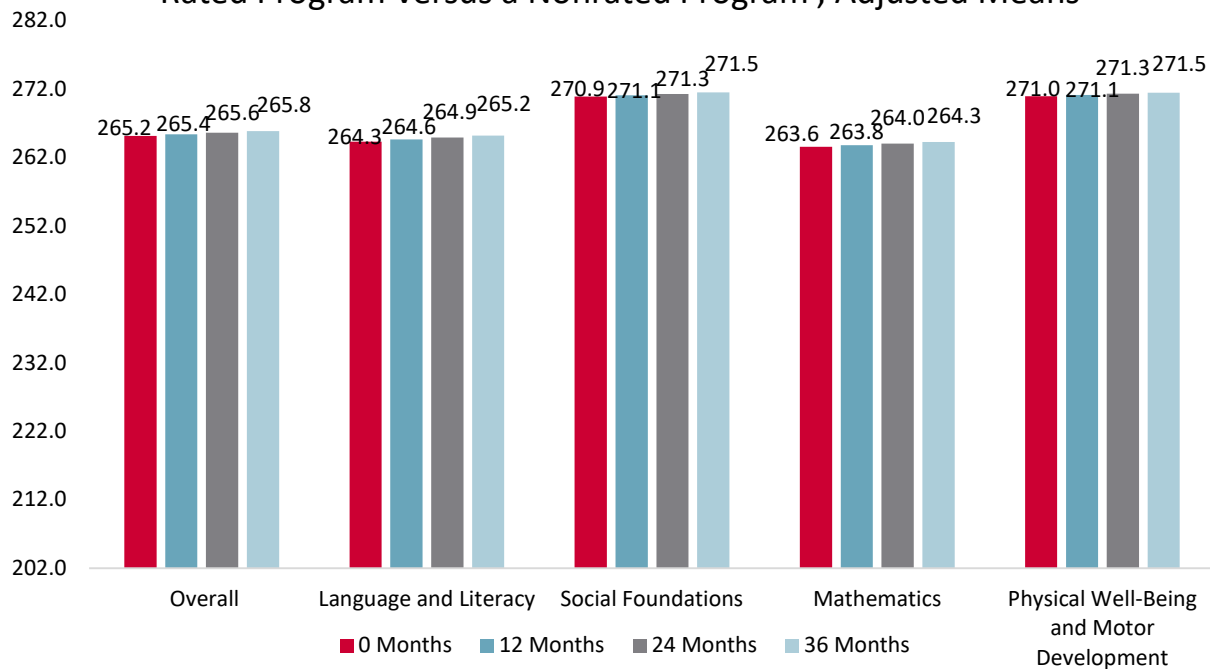
Do Kindergarten Readiness Assessment Scores Vary by Quality Rating?

Comparisons across all Star Rating Levels

After accounting for program and child characteristics, there is significant variability in KRA scores based on quality rating. These are typically considered small effects.[34] In the overall KRA score and all domains, children in 5-Star programs score statistically significantly higher than children in 1-Star programs, with the adjusted mean differences ranging from .96 (Physical Well-Being and Motor Development) to 1.91 (Language and Literacy) points higher.[35] Further, in Language and Literacy, children in 5-Star programs scored statistically significantly higher than children in all other rating levels. In the overall KRA score, Mathematics, and Language and Literacy, children in 4-Star programs also score statistically significantly higher than children in 1-Star programs (and significantly higher than children in 2-Star programs in Language and Literacy).[36] Children in 3-Star programs score statistically significantly higher than children in 1-Star programs in Language and Literacy.[37] Across all domains, there were not statistically significant differences in mean scores between children in 1- and 2-Star, 2- and 3-Star, and 3- and 4-Star programs, indicating that child outcomes are not differentiated based on whether a child is in a 1-Star or 2-Star program, a 2-Star or 3-Star program, or a 3-Star and 4-Star program. Thus, the effects of higher quality on child outcomes are most apparent when comparing the highest and lowest star ratings and the differences are minimal when comparing the middle quality ratings. Comparing scores across all star

---

[34] The standardized mean difference closely resembles that of Cohen's d (a commonly understood effect size) and is interpreted as such. These effects are classified as small according to standards commonly used in the social sciences (i.e., small = .20, medium = .50, large = .80): Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155-159.

[35] The standardized mean difference for Physical Development and Well Being is 0.05 and is 0.15 for Language and Literacy.

[36] Standardized mean differences range from 0.06 for Overall KRA 4- to 1-Star comparison to 0.08 for the Language and Literacy 4- to 1-Star comparison

[37] Standardized mean difference = 0.04

rating levels, scores become increasingly higher as star rating increases in the Overall KRA, Language and Literacy, and Mathematics, though not always statistically significant; this trend was not observed in Social Foundations and Physical Well-Being and Motor Development.

Taken together, children in 4- and 5-Star programs outperform their peers in 1-Star programs in all (5-Star) or most (4-Star) of the KRA domains. The benefits of being in the highest rating level are most apparent in Language and Literacy with children in 5-Star programs scoring statistically significantly higher than their peers in all other rating levels. Lastly, across all domains, child outcomes do not statistically significantly differ between 1-and 2-Star programs, 2- and 3-Star programs, and 3- and 4-Star programs. Figure 4 displays the adjusted mean KRA scores for each of the star ratings and Tables 6 through 10 display mean differences by star rating. The multilevel model statistics are reported in Appendix B (Models 2.1 through 2.5).

Figure 4. KRA Overall and Domain Scores by 5-Star Rating, Adjusted Means
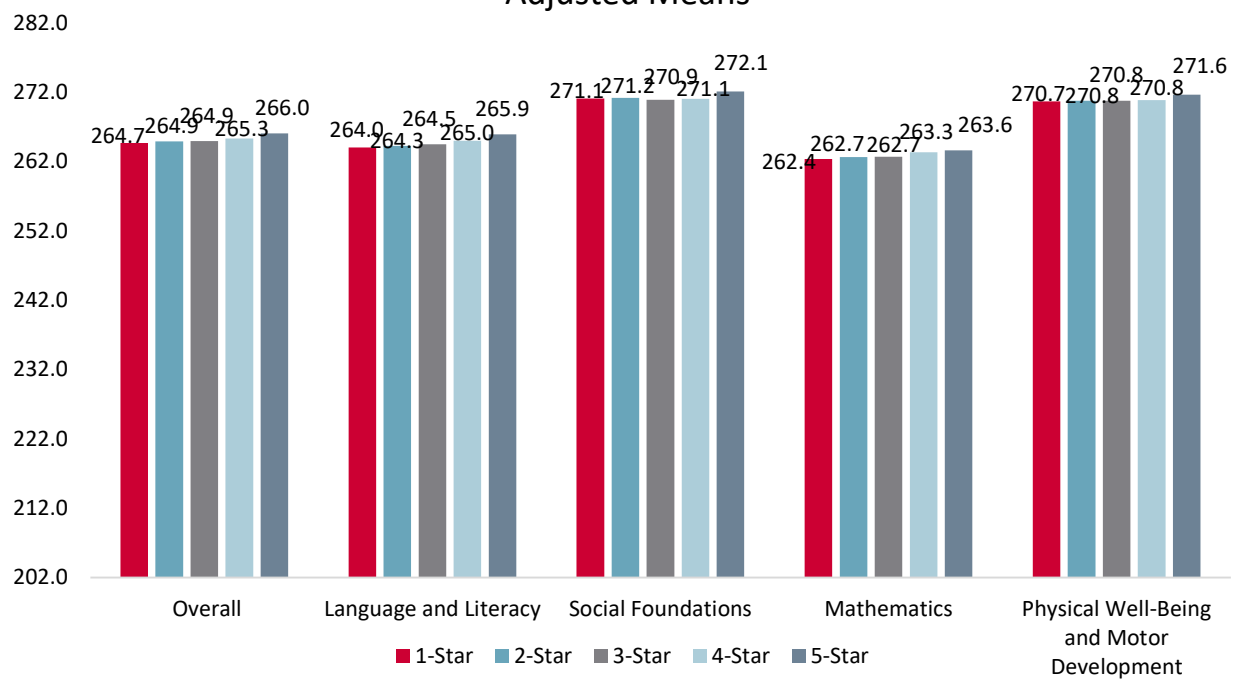
Table 6. Overall KRA Scores, Adjusted Mean Differences				
	4-Star	**3-Star**	**2-Star**	**1-Star**
5-Star	0.71	1.11***	1.14***	1.38***
4-Star		0.40	0.43	0.67**
3-Star			0.03	0.27
2-Star				0.24

*Note. Differences are row means minus column means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Table 7. Language and Literacy, Adjusted Mean Differences	4-Star	3-Star	2-Star	1-Star
5-Star	0.91*	1.42***	1.65***	1.91***
4-Star		0.51	0.74*	1.00***
3-Star			0.23	0.49*
2-Star				0.26

*Note. Differences are row means minus column means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Table 8. Social Foundations, Adjusted Mean Differences	4-Star	3-Star	2-Star	1-Star
5-Star	1.07	1.24*	0.94	1.07*
4-Star		0.17	-0.12	0.00
3-Star			-0.30	-0.17
2-Star				0.13

*Note. Differences are row means minus column means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Table 9. Mathematics, Adjusted Mean Differences	4-Star	3-Star	2-Star	1-Star
5-Star	0.29	0.92**	0.96**	1.22***
4-Star		0.63	0.67	0.93***
3-Star			0.04	0.30
2-Star				0.27

*Note. Differences are row means minus column means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Table 10. Physical Well-Being and Motor Development, Adjusted Mean Differences	4-Star	3-Star	2-Star	1-Star
5-Star	0.81	0.90	0.87	0.96*
4-Star		0.09	0.07	0.15
3-Star			-0.03	0.06
2-Star				0.09

*Note. Differences are row means minus column means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Lower and Higher Star Comparison

Overall, children in higher-rated programs (defined as 3- to 5-Star programs) scored statistically significantly higher in the overall KRA, Language and Literacy, and Mathematics compared to children in lower-rated programs (defined as 1- or 2-Star programs) after accounting for program and child characteristics. These are typically considered small effects.[38] The largest mean difference between higher and lower-rated programs was in Language and Literacy with children in higher-rated programs having, on average, scores that were .81 points higher than those in lower-rated programs. Scores for children in higher-rated programs were .51 and .56 points higher than children in lower-rated programs in the overall KRA and Mathematics, respectively. There were not statistically significant differences between children in higher- and lower-rated programs in Social Foundations or Physical Well-Being and Motor Development. These data indicate that after accounting for child demographics and program characteristics, children who attend highly rated programs have more positive outcomes on the overall KRA and Language and Literacy and Mathematics domains compared to children who attend lower rated

---

[38] Standardized mean difference for Language and Literacy = 0.06; Standardized mean difference for Mathematics = 0.04; Standardized mean difference for Overall KRA = 0.04

programs, though the effects are small. Figure 5 displays the adjusted mean KRA scores for each of the star rating categories and Table 11 displays the mean differences. The multilevel model statistics are reported in Appendix B (Models 2.6 through 2.10).

Figure 5. KRA Overall and Domain Scores by Lower and Higher Star Ratings, Adjusted Means
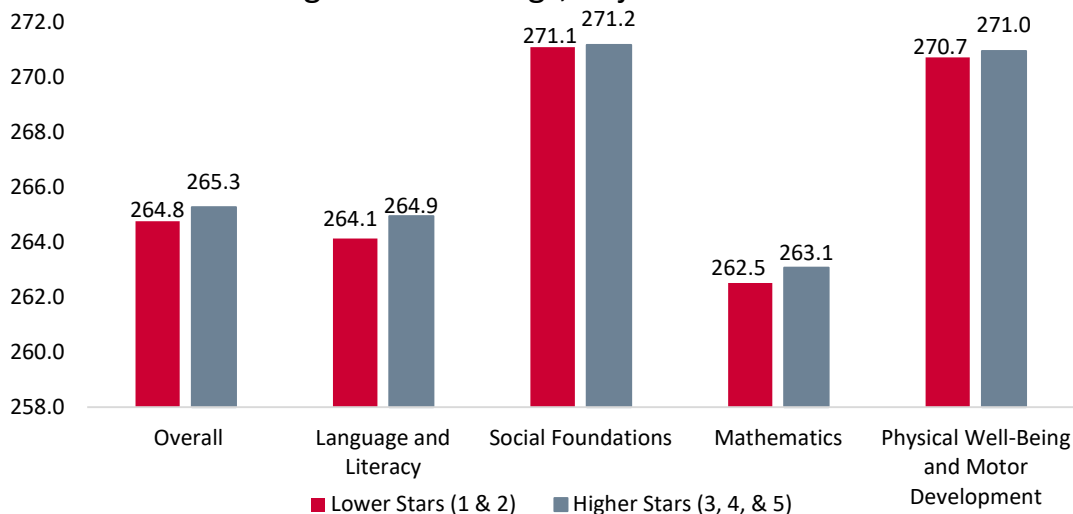
Table 11. KRA Adjusted Mean Differences by Higher- and Lower-Rated Programs	
KRA Domain	**Mean Difference**
KRA Overall	0.51***
Language and Literacy	0.81***
Social Foundations	0.08
Mathematics	0.56***
Physical Well-Being and Motor Development	0.24

*Note. Differences are higher-rated program means minus lower-rated program means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Do Third Grade English Language Arts Test Scores Vary by Quality Rating?

Comparisons across all Star Rating Levels

There is significant variability in third grade ELA scores based on quality rating after accounting for program and child characteristics. Children who attended 3- and 4-Star programs prior to their third grade ELA test administrations scored statistically significantly higher on third grade ELA compared to children in 1-Star programs. These effects are typically considered small.[39] Compared to children in 1-Star programs, children in 3-Star programs scored 2.62 points higher and children in 4-Star programs scored 3.36 points higher. Unlike the KRA scores, children who attended 5-Star programs prior to their third grade ELA test administration did not score statistically significantly higher on third grade ELA when compared to children who attended 1-Star programs. These analyses demonstrate that there are significant differences in third grade ELA scores based on the quality of the early learning program a child attended through third grade, but that the relationship is not linear. Stated another way, the

---

[39] Standardized mean difference for 4-Star compared to 1-Star = 0.09; Standardized mean difference for 3-Star compared to 1-Star = 0.07

scores children achieve on the third grade ELA do not go up consistently as the rating of the programs they attended prior to their third grade ELA test go up. Figure 6 displays the adjusted mean third grade ELA scores for each of the star rating categories and Table 12 displays the mean differences. The multilevel model statistics are reported in Appendix B (Model 3.1).

Figure 6. Third Grade ELA Scores by 5-Star Rating, Adjusted Means
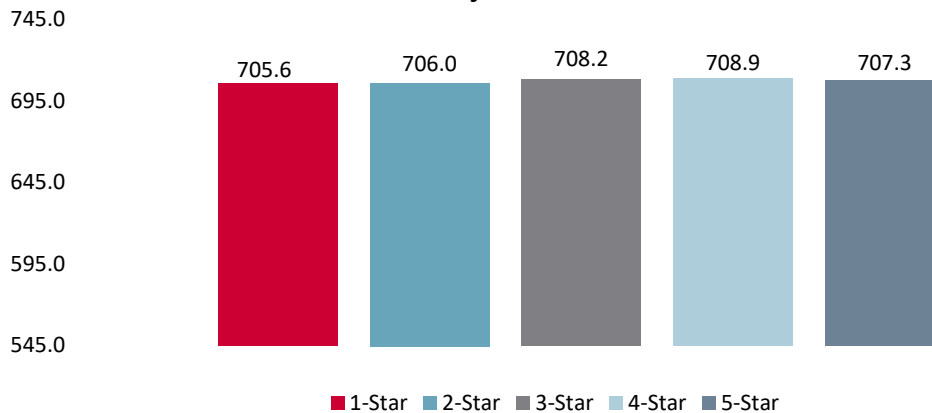
Table 12. Third Grade ELA Adjusted Mean Differences				
	4-Star	**3-Star**	**2-Star**	**1-Star**
5-Star	-1.62	-0.88	1.30	1.74
4-Star		0.73	2.91	3.36**
3-Star			2.18	2.62**
2-Star				0.44

*Note. Differences are higher-rated program means minus lower-rated program means. Significant codes: * = p < .05, ** = p < .01, *** = p < .001*

Lower and Higher Star Comparison

Differences in third grade ELA were also assessed between lower- and higher-rated programs. Overall there were statistically significant differences in third grade ELA scores between higher- and lower-rated programs after accounting for child and program characteristics, with children who attended higher-rated programs scoring, on average, 2.55 points higher than children who attended lower-rated programs (Figure 7); this effect is classified as small.[40] The multilevel model statistics are reported in Appendix B (Model 3.2).

---

[40] Standardized mean difference = 0.07

Figure 7. Third Grade ELA Scores by Lower and Higher Ratings, Adjusted Means
(Mean Difference = 2.55)

(Bar chart showing Third Grade ELA scores. Y-axis ranges from 545.0 to 745.0. Lower Stars (1 & 2) = 705.8; Higher Stars (3, 4, & 5) = 708.3.)

Legend: ■ Lower Stars (1 & 2)   ■ Higher Stars (3, 4, & 5)
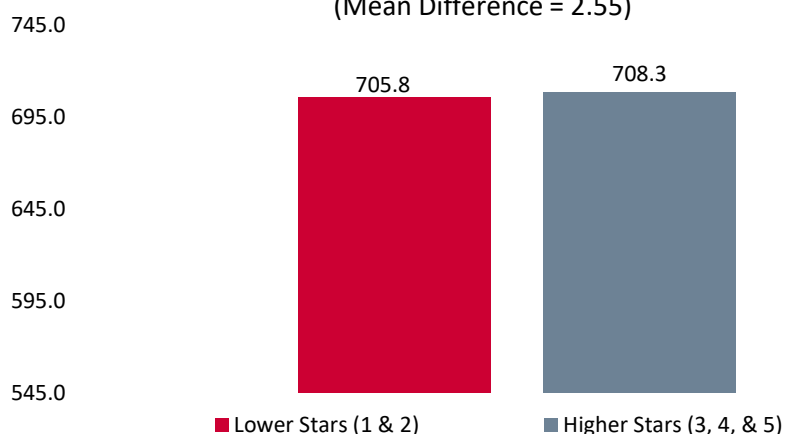
Does a Child's Progress Between Kindergarten and Third Grade Vary by Quality Rating?

A child's progress from KRA to third grade ELA does not vary based on quality rating. After controlling for child and program characteristics, developmental gains between kindergarten and third grade are not dependent upon the quality of the early learning program a child attended. As expected, overall KRA scores strongly predicted third grade ELA scores (for every point increase in KRA scores, third grade ELA scores are predicted to increase by 1.41 points). The relationship between overall KRA and third grade ELA scores did not change based on the quality of early learning program the child attended (i.e. the effect of quality after accounting for KRA was not a statistically significant predictor of third grade ELA across the five levels of quality or in higher to lower rated comparisons). Figures 8 and 9 display adjusted means for each star rating after accounting for KRA scores. The multilevel model statistics are presented in Appendix B (Models 4.1 and 4.2).

Figure 8. Third Grade ELA Scores by 5-Star Rating, Adjusted Means Accounting for KRA

(Bar chart showing Third Grade ELA scores by star rating. Y-axis ranges from 545.0 to 745.0. 1-Star = 703.1; 2-Star = 703.8; 3-Star = 704.3; 4-Star = 705.4; 5-Star = 704.4.)

Legend: ■ 1-Star   ■ 2-Star   ■ 3-Star   ■ 4-Star   ■ 5-Star
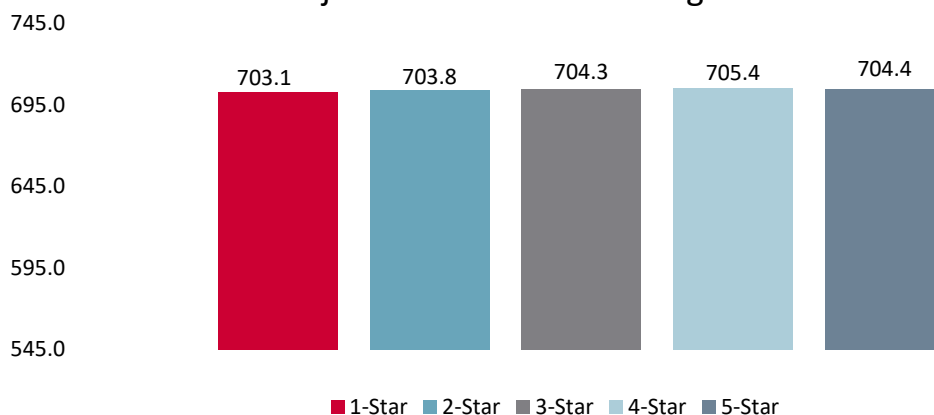
Figure 9. Third Grade ELA Scores by Lower and
Higher Ratings, Adjusted Means Accounting for
KRA

	703.5	704.8

Lower Stars (1 & 2) ■ Higher Stars (3, 4, & 5)
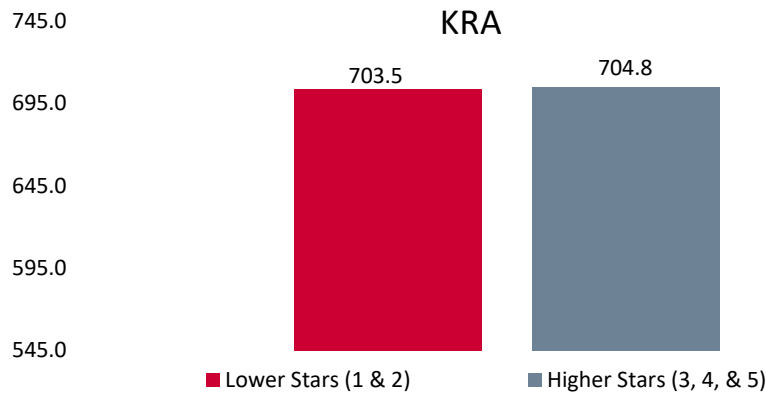
Conclusion

The current study found positive child development outcomes as a result of participating in programs that were rated highly in SUTQ. Specifically, children eligible and receiving publicly funded child care who attended higher-rated programs at any time between birth and their third grade ELA test administration scored higher than children in lower-rated programs in both the KRA and third grade ELA. When we compared children in PFCC, children in rated programs scored statistically significantly higher in all KRA domains than children in non-rated programs. However, similar to previous studies, the levels of program rating were not consistently predictive of child outcomes measured by kindergarten or third grade assessments.[41] In analyses predicting third grade ELA scores, children in 3- and 4-Star programs scored, on average, higher than children in 5-Star programs (though non-significantly) and in KRA analyses children in 3-Star programs did not have statistically significant higher scores than children in 2-Star programs; the same is true in comparing child scores from 3- and 4-Star programs and from 1- and 2-Star programs. These findings indicate that overall quality in early learning programs, as measured by SUTQ, has a positive impact on child outcomes though the trends are, at times, inconsistent.

---

[41] https://www.mathematica.org/our-publications-and-findings/publications/are-ratings-from-tiered-quality-rating-and-improvement-systems-valid-measures-of-program-quality

Appendix A. Variables and Descriptions

Program star ratings over time: Data were provided for each program which included star ratings for each month the program has been in operation since 2012. Only star ratings from October 2013 onwards were included in the analysis as this is the date in which the current 5-Star system of SUTQ began.

Child star ratings over time: Because program star ratings change over time and children often change programs over time, weighted star ratings were computed for each child which provided the proportion of the child's time spent in each star rating during their time in PFCC in Ohio. Weightings were computed separately for the KRA and third-grade ELA analyses. For the KRA analyses, weightings were computed for the time spent in various ratings prior to the KRA administration. For the third grade ELA analyses, weightings were computed for time spent in the various ratings prior to the third grade ELA administration. For example, if a child attended a 2-Star program for 20 percent of their time in PFCC in Ohio prior to their KRA administrations and spent the remaining 80 percent of their pre-KRA time in a 3-Star program, the child's weightings for the KRA analyses for the 1-, 4- and 5-Star programs would be zero and their value for the weighted 2-Star and 3-Star would be .20 and .80, respectively. This method was chosen as it captures the differing proportions of early learning quality, as measured by SUTQ, which children received throughout their enrollment in PFCC in Ohio.

Program zip code and region designation: The addresses and zip codes were provided for each program. Using the zip code, MRC classified programs into either metropolitan, micropolitan, or rural regions using the Rural Urban Commuting Area (RUCA) classification.[42] The RUCA uses measures of population density, urbanization, and daily commuting to classify zip codes on a continuum of rural to urban.

Child enrollment: State student identifiers (SSID) were provided for all children who attended a PFCC program from 2012 to 2019, including the associated program the child was enrolled in each month. Only children enrolled in programs from October 2013 through November 2019 were included in the analysis to align with the current version of SUTQ.

Child demographics: Child family federal poverty level, age at the time each test was taken, gender, race, and household size were provided for each child, when data were available.

Number of months child was enrolled in PFCC in Ohio: Using the child enrollment data, total months enrolled in PFCC was calculated for each child during their time attending early learning programs up until the time of the test administrations (i.e. up until the KRA for the KRA analyses, and up until the third grade ELA for the ELA analyses)? (Same question as above).

Program average federal poverty level (FPL): Using the child FPL, an average FPL was calculated for each program by aggregating all children's FPL (those receiving PFCC) enrolled in a particular program in a given month.

---

[42] https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx

Appendix B. Model Estimates

Model 1.1 KRA Overall between Children in Rated and Unrated Programs

Number of Children: 89,2223 Number of Early Learning Programs: 8,141 Number of Districts: 841

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	2.39	1.55
KRA School District	12.00	3.46
Residual[43]	120.08	10.96

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	265.20	0.38	7,388	693.30	< .0001*
Months Spent in Rated Programs	0.019	0.00	29,070	4.22	< .0001*
Type A (reference = Centers)	-1.34	0.33	19,820	-4.09	< .0001*
Type B (reference = Centers)	-1.16	0.15	27,200	-7.78	< .0001*
ODE (reference = Centers)	0.86	0.67	54,790	1.28	0.20
Program Minority %	-0.02	0.00	18,950	-6.36	< .0001*
Program Average Federal Poverty Level	0.04	0.00	28,270	13.83	< .0001*
Micropolitan (reference = Rural)	0.57	0.38	10,290	1.48	0.14
Metropolitan (reference = Rural)	0.68	0.37	9,315	1.84	0.07
Child months enrolled in Early Learning	0.05	0.00	83,740	22.09	< .0001*
Child Racial Minority	-1.11	0.11	88,690	-10.01	< .0001*
Child Household Size	-0.42	0.03	86,110	-15.14	< .0001*
Child Average Federal Poverty Level	0.01	0.00	88,500	16.97	< .0001*
Year Child took KRA	-0.36	0.03	82,980	-12.96	< .0001*
Child Age at KRA Administration	2.29	0.06	88,620	36.87	< .0001*
Gender (boys compared to girls)	-3.61	0.07	88,340	-48.62	< .0001*

*Note. * indicates statistical significance at p < .05*

Model 1.2 KRA Language and Literacy between Children in Rated and Unrated Programs

Number of Children: 89,636 Number of Early Learning Programs: 8,159 Number of Districts: 842

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	3.20	1.79
KRA School District	9.15	3.03
Residual	138.76	11.78

---

[43] Residual represents unexplained within group variance

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	264.30	0.40	7,184	662.66	< .0001*
Months Spent in Rated Programs	0.024	0.005	31,780	5.06	< .0001*
Type A (reference = Centers)	-1.657	0.36	19,500	-4.67	< .0001*
Type B (reference = Centers)	-1.55	0.16	26,660	-9.63	< .0001*
ODE (reference = Centers)	1.80	0.72	54,910	2.50	< .01*
Program Minority %	-0.01	0.00	17,170	-4.85	< .0001*
Program Average Federal Poverty Level	0.04	0.00	29,240	13.05	< .0001*
Micropolitan (reference = Rural)	0.84	0.41	8,826	2.08	0.04*
Metropolitan (reference = Rural)	0.90	0.39	7,910	2.32	0.02*
Child months enrolled in Early Learning	0.07	0.00	84,950	26.28	< .0001*
Child Racial Minority	-1.26	0.12	89,100	-10.63	< .0001*
Child Household Size	-0.70	0.03	87,060	-23.39	< .0001*
Child Average Federal Poverty Level	0.02	0.00	88,950	17.45	< .0001*
Year Child took KRA	-0.73	0.03	83,950	-24.28	< .0001*
Child Age at KRA Administration	2.37	0.07	89,070	35.60	< .0001*
Gender (boys compared to girls)	-2.45	0.08	88,770	-30.80	< .0001*

*Note. * indicates statistical significance at p < .05*

Model 1.3 KRA Mathematics between Children in Rated and Unrated Programs

Number of Children: 89,691 Number of Early Learning Programs: 8,162 Number of Districts: 842

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	2.08	1.44
KRA School District	9.20	3.03
Residual	142.19	11.92

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	263.60	0.39	6,879	672.32	< .0001*
Months Spent in Rated Programs	0.020	0.00	24,730	4.12	< .0001*
Type A (reference = Centers)	-1.34	0.35	20,930	-3.84	< .0001*
Type B (reference = Centers)	-1.35	0.16	28,640	-8.44	< .0001*
ODE (reference = Centers)	0.67	0.72	56,950	0.92	0.36
Program Minority %	-0.02	0.00	16,320	-6.35	< .0001*
Program Average Federal Poverty Level	0.03	0.00	25,970	10.81	< .0001*
Micropolitan (reference = Rural)	0.42	0.40	8,437	1.07	0.29
Metropolitan (reference = Rural)	0.57	0.38	7,597	1.50	0.13

Child months enrolled in Early Learning	0.05	0.00	82,680	20.24	< .0001*
Child Racial Minority	-1.68	0.12	89,390	-14.04	< .0001*
Child Household Size	-0.44	0.03	85,460	-14.51	< .0001*
Child Average Federal Poverty Level	0.01	0.00	89,180	15.67	< .0001*
Year Child took KRA	-0.97	0.03	82,090	-32.15	< .0001*
Child Age at KRA Administration	2.76	0.07	89,280	41.06	< .0001*
Gender (boys compared to girls)	-1.43	0.08	89,010	-17.80	< .0001*

*Note. * indicates statistical significance at p < .05*

Model 1.4 KRA Social Foundations between Children in Rated and Unrated Programs

Number of Children: 89,445 Number of Early Learning Programs: 8,153 Number of Districts: 841

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	3.33	1.83
KRA School District	42.65	6.53
Residual	363.36	19.06

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	270.90	0.65	7,340	415.55	< .0001*
Months Spent in Rated Programs	0.017	0.01	21,560	2.22	0.03*
Type A (reference = Centers)	-1.33	0.55	24,560	-2.41	0.02*
Type B (reference = Centers)	-0.39	0.25	32,490	-1.54	0.12
ODE (reference = Centers)	0.46	1.15	60,790	0.40	0.69
Program Minority %	-0.02	0.00	19,710	-4.41	< .0001*
Program Average Federal Poverty Level	0.06	0.00	25,290	12.06	< .0001*
Micropolitan (reference = Rural)	1.06	0.64	11,530	1.64	0.10
Metropolitan (reference = Rural)	1.08	0.63	10,510	1.73	0.08
Child months enrolled in Early Learning	0.05	0.00	80,670	11.17	< .0001*
Child Racial Minority	-0.96	0.19	89,160	-5.01	< .0001*
Child Household Size	-0.25	0.05	84,030	-5.11	< .0001*
Child Average Federal Poverty Level	0.01	0.00	88,890	10.27	< .0001*
Year Child took KRA	0.85	0.05	80,810	17.67	< .0001*
Child Age at KRA Administration	1.94	0.11	88,990	18.04	< .0001*
Gender (boys compared to girls)	-7.36	0.13	88,780	-57.29	< .0001*

Model 1.5 KRA Physical Well-Being and Development between Children in Rated and Unrated Programs

Number of Children: 89,439 Number of Early Learning Programs: 8,150 Number of Districts: 841

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	2.69	1.64
KRA School District	33.25	5.77
Residual	279.96	16.73

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	271.00	0.57	7,309	472.21	< .0001*
Months Spent in Rated Programs	0.014	0.01	21,990	2.15	0.03*
Type A (reference = Centers)	-1.34	0.49	24,410	-2.76	0.01*
Type B (reference = Centers)	-0.98	0.22	32,360	-4.39	< .0001*
ODE (reference = Centers)	-0.33	1.01	60,680	-0.33	0.74
Program Minority %	-0.02	0.00	19,800	-5.40	< .0001*
Program Average Federal Poverty Level	0.05	0.00	25,530	11.16	< .0001*
Micropolitan (reference = Rural)	0.55	0.57	11,570	0.98	0.33
Metropolitan (reference = Rural)	0.74	0.55	10,540	1.34	0.18
Child months enrolled in Early Learning	0.05	0.00	80,870	14.40	< .0001*
Child Racial Minority	0.18	0.17	89,140	1.04	0.30
Child Household Size	-0.04	0.04	84,220	-1.02	0.31
Child Average Federal Poverty Level	0.01	0.00	88,870	9.68	< .0001*
Year Child took KRA	0.26	0.04	81,040	6.19	< .0001*
Child Age at KRA Administration	2.00	0.09	88,980	21.21	< .0001*
Gender (boys compared to girls)	-7.91	0.11	88,760	-70.08	< .0001*

*Note. * indicates statistical significance at p < .05*

Model 2.1 Overall KRA Scores, 5-Star Comparison

Number of Children: 45,536 Number of Early Learning Programs: 3,787 Number of Districts: 823

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	2.53	1.59
KRA School District	12.77	3.57
Residual	117.91	10.86

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	264.70	0.57	4,149	468.052	< .0001*
Weighted 2-Star (reference = 1-Star)	0.24	0.20	4,124	1.224	0.22
Weighted 3-Star (reference = 1-Star)	0.27	0.19	3,965	1.418	0.16
Weighted 4-Star (reference = 1-Star)	0.67	0.24	4,820	2.845	< 0.01*
Weighted 5-Star (reference = 1-Star)	1.38	0.26	5,311	5.296	< .0001*
Type A (reference = Centers)	-1.88	0.53	17,810	-3.577	< .0001*
Type B (reference = Centers)	-1.51	0.26	30,720	-5.725	< .0001*
ODE (reference = Centers)	-0.58	1.12	42,120	-0.516	0.61
Program Minority %	-0.02	0.00	8,722	-4.252	< .0001*
Program Average Federal Poverty Level	0.04	0.00	12,720	9.94	< .0001*
Micropolitan (reference = Rural)	1.25	0.58	4,213	2.15	0.03*
Metropolitan (reference = Rural)	1.46	0.55	4,060	2.647	< 0.01*
Child months enrolled in Early Learning	0.06	0.00	45,110	19.315	< .0001*
Child Racial Minority	-1.02	0.15	45,210	-6.806	< .0001*
Child Household Size	-0.36	0.04	45,290	-8.729	< .0001*
Child Average Federal Poverty Level	0.01	0.00	45,080	12.46	< .0001*
Year Child took KRA	-0.53	0.04	38,170	-13.064	< .0001*
Child Age at KRA Administration	2.48	0.09	45,140	27.794	< .0001*
Gender (boys compared to girls)	-3.67	0.10	44,930	-35.559	< .0001*

*Note. * indicates statistical significance at p < .05*

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	0.71	0.30	2.40	0.08
5-Star to 3-Star	1.11	0.27	4.11	< .001*
5-Star to 2-Star	1.14	0.28	4.10	< 0.001*
4-Star to 3-Star	0.40	0.24	1.64	0.36
4-Star to 2-Star	0.43	0.25	1.70	0.32
3-Star to 2-Star	0.03	0.21	0.15	1.00

Model 2.2 KRA Language and Literacy Scores, 5-Star Comparison

Number of Children: 45,679 Number of Early Learning Programs: 3,786 Number of Districts: 824

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	3.02	1.74
KRA School District	10.73	3.28
Residual	135.91	11.66

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	264.00	0.59	3,888	446.37	< .0001*
Weighted 2-Star (reference = 1-Star)	0.26	0.21	4,333	1.25	0.21
Weighted 3-Star (reference = 1-Star)	0.49	0.21	4,164	2.40	0.02*
Weighted 4-Star (reference = 1-Star)	1.00	0.25	5,077	3.97	< .0001*
Weighted 5-Star (reference = 1-Star)	1.91	0.28	5,573	6.83	< .0001*
Type A (reference = Centers)	-1.95	0.56	18,100	-3.47	< .0001*
Type B (reference = Centers)	-1.83	0.28	30,920	-6.48	< .0001*
ODE (reference = Centers)	0.10	1.20	42,380	0.08	0.93
Program Minority %	-0.01	0.00	8,208	-2.72	0.01*
Program Average Federal Poverty Level	0.04	0.00	13,060	9.00	< .0001*
Micropolitan (reference = Rural)	1.40	0.61	3,821	2.30	0.02*
Metropolitan (reference = Rural)	1.39	0.58	3,668	2.41	0.02*
Child months enrolled in Early Learning	0.07	0.00	45,350	23.23	< .0001*
Child Racial Minority	-1.21	0.16	45,410	-7.51	< .0001*
Child Household Size	-0.63	0.04	45,510	-14.50	< .0001*
Child Average Federal Poverty Level	0.02	0.00	45,280	13.87	< .0001*
Year Child took KRA	-0.89	0.04	38,570	-20.39	< .0001*
Child Age at KRA Administration	2.49	0.10	45,340	26.10	< .0001*
Gender (boys compared to girls)	-2.46	0.11	45,130	-22.28	< .0001*

*Note. * indicates statistical significance at p < .05*

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	0.91	0.32	2.86	0.02*
5-Star to 3-Star	1.42	0.29	4.90	<.001*
5-Star to 2-Star	1.65	0.30	5.52	<.001*
4-Star to 3-Star	0.51	0.26	1.95	0.20
4-Star to 2-Star	0.74	0.27	2.73	0.03*
3-Star to 2-Star	0.23	0.23	1.01	0.74

Model 2.3 KRA Mathematics Scores, 5-Star Comparison

Number of Children: 45,718 Number of Early Learning Programs: 3,789 Number of Districts: 824

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	2.05	1.43
KRA School District	10.54	3.25
Residual	136.66	11.69

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	262.40	0.58	3,724	455.92	< .0001*
Weighted 2-Star (reference = 1-Star)	0.27	0.21	3,607	1.31	0.19
Weighted 3-Star (reference = 1-Star)	0.30	0.20	3,551	1.52	0.13
Weighted 4-Star (reference = 1-Star)	0.93	0.25	4,238	3.81	< 0.001*
Weighted 5-Star (reference = 1-Star)	1.22	0.27	4,603	4.49	< .0001*
Type A (reference = Centers)	-1.80	0.56	18,870	-3.23	< 0.01*
Type B (reference = Centers)	-1.65	0.28	32,070	-5.87	< .0001*
ODE (reference = Centers)	-1.30	1.20	42,740	-1.09	0.28
Program Minority %	-0.02	0.00	7,725	-4.74	< .0001*
Program Average Federal Poverty Level	0.04	0.00	10,990	8.47	< .0001*
Micropolitan (reference = Rural)	1.71	0.59	3,633	2.89	<.01*
Metropolitan (reference = Rural)	1.83	0.56	3,506	3.25	<.01*
Child months enrolled in Early Learning	0.06	0.00	45,150	17.98	< .0001*
Child Racial Minority	-1.72	0.16	45,530	-10.66	< .0001*
Child Household Size	-0.40	0.04	45,460	-9.13	< .0001*
Child Average Federal Poverty Level	0.01	0.00	45,400	11.41	< .0001*
Year Child took KRA	-1.04	0.04	37,110	-23.88	< .0001*
Child Age at KRA Administration	2.84	0.10	45,450	29.66	< .0001*
Gender (boys compared to girls)	-1.44	0.11	45,250	-12.98	< .0001*

Note. * indicates statistical significance at p < .05

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	0.29	0.31	0.93	0.79
5-Star to 3-Star	0.92	0.28	3.26	0.01*
5-Star to 2-Star	0.96	0.29	3.30	0.01*
4-Star to 3-Star	0.63	0.25	2.48	0.06
4-Star to 2-Star	0.67	0.26	2.54	0.05
3-Star to 2-Star	0.04	0.22	0.16	1.00

Model 2.4 KRA Physical Well-Being and Development Scores, 5-Star Comparison

Number of Children: 45,614 Number of Early Learning Programs: 3,788 Number of Districts: 823

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	3.14	1.77
KRA School District	33.52	5.79
Residual	278.43	16.69

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	270.70	0.84	4,230	320.47	< .0001*
Weighted 2-Star (reference = 1-Star)	0.09	0.29	3,354	0.31	0.76
Weighted 3-Star (reference = 1-Star)	0.06	0.28	3,371	0.22	0.82
Weighted 4-Star (reference = 1-Star)	0.15	0.34	3,937	0.45	0.65
Weighted 5-Star (reference = 1-Star)	0.96	0.38	4,228	2.50	0.01*
Type A (reference = Centers)	-2.05	0.79	20,170	-2.60	0.01*
Type B (reference = Centers)	-1.63	0.40	33,240	-4.08	< .0001*
ODE (reference = Centers)	-1.41	1.71	43,100	-0.83	0.41
Program Minority %	-0.02	0.01	8,590	-3.39	< 0.01*
Program Average Federal Poverty Level	0.05	0.01	10,250	7.30	< .0001*
Micropolitan (reference = Rural)	1.16	0.86	4,412	1.34	0.18
Metropolitan (reference = Rural)	1.48	0.82	4,246	1.80	0.07
Child months enrolled in Early Learning	0.06	0.00	44,760	12.28	< .0001*
Child Racial Minority	0.34	0.23	45,380	1.49	0.14
Child Household Size	0.05	0.06	45,190	0.86	0.39
Child Average Federal Poverty Level	0.01	0.00	45,220	7.37	< .0001*
Year Child took KRA	0.05	0.06	36,360	0.79	0.43
Child Age at KRA Administration	2.29	0.14	45,280	16.77	< .0001*
Gender (boys compared to girls)	-7.91	0.16	45,100	-50.05	< .0001*

*Note. * indicates statistical significance at p < .05*

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	0.81	0.44	1.82	0.26
5-Star to 3-Star	0.90	0.40	2.25	0.11
5-Star to 2-Star	0.87	0.41	2.14	0.14
4-Star to 3-Star	0.09	0.36	0.26	0.99
4-Star to 2-Star	0.07	0.37	0.18	1.00
3-Star to 2-Star	-0.03	0.31	-0.09	1.00

Model 2.5 KRA Social Foundations Scores, 5-Star Comparison

Number of Children: 45,617 Number of Early Learning Programs: 3,789 Number of Districts: 823

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	3.99	2.00
KRA School District	43.80	6.62
Residual	362.70	19.05

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	271.10	0.96	4,216	281.36	< .0001*
Weighted 2-Star (reference = 1-Star)	0.13	0.33	3,237	0.38	0.70
Weighted 3-Star (reference = 1-Star)	-0.17	0.32	3,258	-0.53	0.59
Weighted 4-Star (reference = 1-Star)	0.00	0.39	3,797	0.00	1.00
Weighted 5-Star (reference = 1-Star)	1.07	0.44	4,078	2.45	0.01*
Type A (reference = Centers)	-2.37	0.90	19,920	-2.63	0.01*
Type B (reference = Centers)	-1.18	0.46	33,070	-2.59	0.01*
ODE (reference = Centers)	-0.88	1.96	43,020	-0.45	0.65
Program Minority %	-0.02	0.01	8,456	-3.11	< .0001*
Program Average Federal Poverty Level	0.06	0.01	9,929	8.16	< .0001*
Micropolitan (reference = Rural)	1.38	0.98	4,372	1.41	0.16
Metropolitan (reference = Rural)	1.86	0.94	4,227	1.98	0.05*
Child months enrolled in Early Learning	0.05	0.01	44,730	9.70	< .0001*
Child Racial Minority	-0.74	0.26	45,390	-2.80	0.01*
Child Household Size	-0.13	0.07	45,170	-1.82	0.07
Child Average Federal Poverty Level	0.01	0.00	45,230	6.68	< .0001*
Year Child took KRA	0.51	0.07	36,020	7.24	< .0001*
Child Age at KRA Administration	2.29	0.16	45,280	14.72	< .0001*
Gender (boys compared to girls)	-7.57	0.18	45,110	-41.96	< .0001*

*Note. * indicates statistical significance at p < .05*

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	1.07	0.50	2.12	0.15
5-Star to 3-Star	1.24	0.45	2.73	0.03*
5-Star to 2-Star	0.94	0.46	2.04	0.17
4-Star to 3-Star	0.17	0.41	0.42	0.98
4-Star to 2-Star	-0.12	0.42	-0.30	0.99
3-Star to 2-Star	-0.30	0.35	-0.84	0.83

Model 2.6 Overall KRA Scores, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 45,536 Number of Early Learning Programs: 3,787 Number of Districts: 823

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	2.54	1.60
KRA School District	12.76	3.57
Residual	117.96	10.86

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	264.80	0.56	4,035	473.97	< .0001*
Weighted Higher Star (reference = Lower Star)	0.51	0.15	3,104	3.49	< .0001*
Type A (reference = Centers)	-1.92	0.53	17,740	-3.66	< .0001*
Type B (reference = Centers)	-1.53	0.26	30,570	-5.84	< .0001*
ODE (reference = Centers)	-0.45	1.12	42,130	-0.40	0.69
Program Minority %	-0.02	0.00	8,631	-4.05	< .0001*
Program Average Federal Poverty Level	0.04	0.00	12,710	9.91	< .0001*
Micropolitan (reference = Rural)	1.27	0.58	4,209	2.18	0.03*
Metropolitan (reference = Rural)	1.44	0.55	4,047	2.61	0.01 *
Child months enrolled in Early Learning	0.06	0.00	45,080	19.21	< .0001*
Child Racial Minority	-1.03	0.15	45,210	-6.84	< .0001*
Child Household Size	-0.36	0.04	45,290	-8.73	< .0001*
Child Average Federal Poverty Level	0.01	0.00	45,080	12.52	< .0001*
Year Child took KRA	-0.50	0.04	40,870	-12.59	< .0001*
Child Age at KRA Administration	2.46	0.09	45,120	27.63	< .0001*
Gender (boys compared to girls)	-3.67	0.10	44,930	-35.53	< .0001*

Note. * indicates statistical significance at p < .05

Model 2.7 KRA Language and Literacy Scores, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 45,679 Number of Early Learning Programs: 3,786 Number of Districts: 824

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	3.02	1.74
KRA School District	10.72	3.28
Residual	136.00	11.66

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	264.10	0.58	3,769	452.47	< .0001*
Weighted Higher Star (reference = Lower Star)	0.81	0.16	3,256	5.20	< .0001*
Type A (reference = Centers)	-2.00	0.56	18,040	-3.56	< .0001*
Type B (reference = Centers)	-1.85	0.28	30,780	-6.59	< .0001*
ODE (reference = Centers)	0.26	1.20	42,390	0.22	0.83
Program Minority %	-0.01	0.00	8,132	-2.49	0.01*
Program Average Federal Poverty Level	0.04	0.00	13,010	8.97	<.0001*

Micropolitan (reference = Rural)	1.43	0.61	3,816	2.34	0.02*
Metropolitan (reference = Rural)	1.36	0.58	3,657	2.36	0.02 *
Child months enrolled in Early Learning	0.07	0.00	45,320	23.11	< .0001*
Child Racial Minority	-1.22	0.16	45,410	-7.55	< .0001*
Child Household Size	-0.63	0.04	45,510	-14.49	< .0001*
Child Average Federal Poverty Level	0.02	0.00	45,290	13.93	< .0001*
Year Child took KRA	-0.84	0.04	41,180	-19.96	< .0001*
Child Age at KRA Administration	2.47	0.10	45,330	25.90	< .0001*
Gender (boys compared to girls)	-2.46	0.11	45,130	-22.26	< .0001*

*Note. * indicates statistical significance at p < .05*

Model 2.8 KRA Mathematics Scores, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 45,718 Number of Early Learning Programs: 3,789 Number of Districts: 824

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	2.05	1.43
KRA School District	10.53	3.24
Residual	136.71	11.69

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	262.50	0.57	3,620	462.35	< .0001*
Weighted Higher Star (reference = Lower Star)	0.56	0.15	2,742	3.69	< .0001*
Type A (reference = Centers)	-1.85	0.56	18,900	-3.32	< .0001*
Type B (reference = Centers)	-1.68	0.28	31,970	-6.00	< .0001*
ODE (reference = Centers)	-1.19	1.20	42,790	-1.00	0.32
Program Minority %	-0.02	0.00	7,664	-4.54	< .0001*
Program Average Federal Poverty Level	0.04	0.00	11,010	8.44	< .0001*
Micropolitan (reference = Rural)	1.72	0.59	3,642	2.92	< .0001*
Metropolitan (reference = Rural)	1.79	0.56	3,505	3.19	< .0001*
Child months enrolled in Early Learning	0.06	0.00	45,110	17.89	< .0001*
Child Racial Minority	-1.73	0.16	45,530	-10.69	< .0001*
Child Household Size	-0.40	0.04	45,460	-9.14	< .0001*
Child Average Federal Poverty Level	0.01	0.00	45,400	11.46	< .0001*
Year Child took KRA	-1.01	0.04	40,160	-23.92	< .0001*
Child Age at KRA Administration	2.82	0.10	45,440	29.54	< .0001*
Gender (boys compared to girls)	-1.43	0.11	45,250	-12.97	< .0001*

Model 2.9 KRA Physical Well-Being and Development Scores, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 45,614 Number of Early Learning Programs: 3,788 Number of Districts: 823

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	3.16	1.78
KRA School District	33.50	5.79
Residual	278.45	16.69

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	270.70	0.83	4,123	324.46	<.0001*
Weighted Higher Star (reference = Lower Star)	0.24	0.21	2,568	1.12	0.26
Type A (reference = Centers)	-2.08	0.79	20,130	-2.63	0.01*
Type B (reference = Centers)	-1.64	0.40	33,080	-4.11	<.0001*
ODE (reference = Centers)	-1.31	1.71	43,130	-0.76	0.45
Program Minority %	-0.02	0.01	8,478	-3.30	<.0001*
Program Average Federal Poverty Level	0.05	0.01	10,270	7.27	<.0001*
Micropolitan (reference = Rural)	1.18	0.86	4,417	1.36	0.17
Metropolitan (reference = Rural)	1.48	0.82	4,241	1.80	0.07
Child months enrolled in Early Learning	0.06	0.00	44,700	12.23	<.0001*
Child Racial Minority	0.34	0.23	45,380	1.48	0.14
Child Household Size	0.05	0.06	45,200	0.87	0.39
Child Average Federal Poverty Level	0.01	0.00	45,220	7.40	<.0001*
Year Child took KRA	0.08	0.06	39,570	1.29	0.20
Child Age at KRA Administration	2.28	0.14	45,270	16.69	<.0001*
Gender (boys compared to girls)	-7.91	0.16	45,100	-50.04	<.0001*

*Note. * indicates statistical significance at p < .05*

Model 2.10 KRA Social Foundations Scores, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 45,617 Number of Early Learning Programs: 3,789 Number of Districts: 823

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	4.02	2.01
KRA School District	43.79	6.62
Residual	362.75	19.05
Fixed Effects		

Variable	Estimate	Standard Error	DF	t value	p value
Intercept	271.10	0.95	4,102	284.85	<.0001*
Weighted Higher Star (reference = Lower Star)	0.08	0.24	2,469	0.34	0.74
Type A (reference = Centers)	-2.41	0.90	19,810	-2.68	0.01*
Type B (reference = Centers)	-1.19	0.46	32,860	-2.62	0.01*
ODE (reference = Centers)	-0.73	1.96	43,050	-0.38	0.71
Program Minority %	-0.02	0.01	8,316	-2.99	<.0001*
Program Average Federal Poverty Level	0.06	0.01	9,910	8.12	<.0001*
Micropolitan (reference = Rural)	1.41	0.98	4,365	1.43	0.15
Metropolitan (reference = Rural)	1.86	0.94	4,212	1.98	0.05
Child months enrolled in Early Learning	0.05	0.01	44,660	9.63	<.0001*
Child Racial Minority	-0.74	0.26	45,390	-2.82	<.0001*
Child Household Size	-0.13	0.07	45,180	-1.81	0.07
Child Average Federal Poverty Level	0.01	0.00	45,230	6.71	<.0001*
Year Child took KRA	0.55	0.07	39,280	8.04	<.0001*
Child Age at KRA Administration	2.28	0.16	45,270	14.61	<.0001*
Gender (boys compared to girls)	-7.56	0.18	45,110	-41.95	<.0001*

*Note. * indicates statistical significance at p < .05*

Model 3.1 Third Grade ELA 5-Star Comparison

Number of Children: 15,509 Number of Early Learning Programs: 2,688 Number of Districts: 731

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	21.28	4.61
Third Grade School District	71.06	8.43
Residual	1,283.15	35.82

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	705.60	2.66	2,811	265.76	<.0001*
Weighted 2-Star (reference = 1-Star)	0.44	1.04	2,603	0.43	0.67
Weighted 3-Star (reference = 1-Star)	2.62	0.95	2,342	2.77	0.01*
Weighted 4-Star (reference = 1-Star)	3.36	1.21	2,314	2.78	0.01*
Weighted 5-Star (reference = 1-Star)	1.74	1.67	3,098	1.04	0.30
Type A (reference = Centers)	-10.58	3.14	11,590	-3.37	<.0001*
Type B (reference = Centers)	-2.22	1.39	13,100	-1.60	0.11
ODE (reference = Centers)	3.71	4.87	15,290	0.76	0.45
Program Minority %	-0.12	0.02	4,419	-6.39	<.0001*

Program Average Federal Poverty Level	0.15	0.02	6,680	6.18	<.0001*
Micropolitan (reference = Rural)	1.24	2.69	2,363	0.46	0.64
Metropolitan (reference = Rural)	3.18	2.54	2,322	1.25	0.21
Child months enrolled in Early Learning	-0.05	0.01	14,740	-3.28	<.0001*
Child Racial Minority	-6.10	0.88	15,410	-6.92	<.0001*
Child Household Size	-1.23	0.23	15,420	-5.24	<.0001*
Child Average Federal Poverty Level	0.04	0.01	15,400	6.53	<.0001*
Year Child took Third Grade ELA	0.04	0.46	15,390	0.09	0.93
Age at Third Grade ELA Administration	0.44	0.59	15,420	0.75	0.45
Gender (boys compared to girls)	-6.84	0.59	15,370	-11.68	<.0001*
Number of districts from K to 3rd	-2.27	0.37	15,460	-6.22	<.0001*

Note. * indicates statistical significance at p < .05

Comparison	Estimate	Standard Error	z value	p value
5-Star to 4-Star	-1.62	1.89	-0.86	0.82
5-Star to 3-Star	-0.88	1.71	-0.52	0.95
5-Star to 2-Star	1.30	1.76	0.74	0.88
4-Star to 3-Star	0.73	1.25	0.59	0.93
4-Star to 2-Star	2.91	1.32	2.20	0.12
3-Star to 2-Star	2.18	1.09	2.00	0.18

Model 3.2 Third Grade ELA, Lower Star (1 and 2-Star) Compared to Higher Star (3, 4, and 5-Star)

Number of Children: 15,509 Number of Early Learning Programs: 2,688 Number of Districts: 731

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	21.32	4.62
Third Grade School District	71.15	8.44
Residual	1,283.18	35.82

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	705.80	2.62	2,696	269.659	<.0001*
Weighted Higher Star (reference = Lower Star)	2.55	0.74	1,883	3.445	<.0001*
Type A (reference = Centers)	-10.61	3.14	11,590	-3.382	<.0001*
Type B (reference = Centers)	-2.29	1.38	13,010	-1.654	0.10
ODE (reference = Centers)	3.70	4.87	15,280	0.761	0.45
Program Minority %	-0.12	0.02	4,391	-6.388	<.0001*
Program Average Federal Poverty Level	0.15	0.02	6,683	6.17	<.0001*

Micropolitan (reference = Rural)	1.25	2.69	2,365	0.466	0.64
Metropolitan (reference = Rural)	3.12	2.54	2,317	1.228	0.22
Child months enrolled in Early Learning	-0.05	0.01	14,810	-3.351	<.0001*
Child Racial Minority	-6.10	0.88	15,410	-6.925	<.0001*
Child Household Size	-1.23	0.23	15,420	-5.259	<.0001*
Child Average Federal Poverty Level	0.04	0.01	15,400	6.552	<.0001*
Year Child took Third Grade ELA	0.03	0.46	15,400	0.056	0.96
Age at Third Grade ELA Administration	0.45	0.59	15,420	0.764	0.45
Gender (boys compared to girls)	-6.84	0.59	15,370	-11.682	<.0001*
Number of districts from K to 3rd	-2.28	0.37	15,460	-6.226	<.0001*

*Note. * indicates statistical significance at p < .05*

Model 4.1 KRA Scores Predicting Third Grade ELA Scores, 5-Star Comparison (i.e. Growth)

Number of Children: 12,176 Number of Early Learning Programs: 2,451 Number of Districts: 718

Random Effects		
Groups	Variance	Standard Deviation
Early Learning Program	8.83	2.97
Third Grade School District	76.66	8.76
Residual	1,080.66	32.87

Fixed Effects					
Variable	Estimate	Standard Error	DF	t value	p value
Intercept	703.10	2.63	2,372	267.90	<.0001*
Weighted 2-Star (reference = 1-Star)	0.68	1.06	2,177	0.64	0.52
Weighted 3-Star (reference = 1-Star)	1.23	0.95	2,108	1.30	0.20
Weighted 4-Star (reference = 1-Star)	2.29	1.21	1,944	1.89	0.06
Weighted 5-Star (reference = 1-Star)	1.31	1.69	2,228	0.78	0.44
Type A (reference = Centers)	-3.67	3.28	10,360	-1.12	0.26
Type B (reference = Centers)	0.66	1.43	10,780	0.46	0.65
ODE (reference = Centers)	3.66	5.47	12,130	0.67	0.50
Program Minority %	-0.11	0.02	3,830	-5.47	<.0001*
Program Average Federal Poverty Level	0.10	0.02	5,065	4.18	<.0001*
Micropolitan (reference = Rural)	-0.78	2.63	1,969	-0.30	0.77
Metropolitan (reference = Rural)	2.52	2.49	1,930	1.01	0.31
Child months enrolled in Early Learning	-0.07	0.01	11,440	-4.81	<.0001*
Child Racial Minority	-4.87	0.90	12,100	-5.38	<.0001*
Child Household Size	-0.54	0.24	12,040	-2.20	0.03*
Child Average Federal Poverty Level	0.03	0.01	12,060	3.97	<.0001*
Year Child took Third Grade ELA	1.34	0.63	12,080	2.11	0.03*
Age at Third Grade ELA Administration	0.04	0.60	12,050	0.06	0.95

Gender (boys compared to girls)	-2.17	0.61	12,030	-3.54	<.0001*
Number of districts from K to 3rd	-1.26	0.38	12,090	-3.28	<.0001*
KRA Overall Score	1.41	0.03	12,120	46.74	<.0001*

*Note. * indicates statistical significance at p < .05*

Model 4.2 KRA Scores Predicting Third Grade ELA Scores, Lower and Higher Star Comparison (i.e. Growth)

Number of Children: 12,176 Number of Early Learning Programs: 2,451 Number of Districts: 718

Random Effects		
Groups	**Variance**	**Standard Deviation**
Early Learning Program	9.06	3.01
Third Grade School District	76.61	8.75
Residual	1,080.58	32.87

Fixed Effects					
Variable	**Estimate**	**Standard Error**	**DF**	**t value**	**p value**
Intercept	703.50	2.58	2,249	272.57	<.0001*
Weighted Higher Star (reference = Lower Star)	1.29	0.74	1,623	1.75	0.08
Type A (reference = Centers)	-3.72	3.28	10,350	-1.13	0.26
Type B (reference = Centers)	0.56	1.42	10,710	0.39	0.70
ODE (reference = Centers)	3.64	5.46	12,130	0.67	0.50
Program Minority %	-0.11	0.02	3,816	-5.43	<.0001*
Program Average Federal Poverty Level	0.10	0.02	5,083	4.16	<.0001*
Micropolitan (reference = Rural)	-0.76	2.63	1,970	-0.29	0.77
Metropolitan (reference = Rural)	2.42	2.49	1,921	0.97	0.33
Child months enrolled in Early Learning	-0.07	0.01	11,450	-4.86	<.0001*
Child Racial Minority	-4.87	0.90	12,100	-5.39	<.0001*
Child Household Size	-0.54	0.24	12,040	-2.23	0.03*
Child Average Federal Poverty Level	0.03	0.01	12,060	3.98	<.0001*
Year Child took Third Grade ELA	1.33	0.63	12,090	2.10	0.04*
Age at Third Grade ELA Administration	0.04	0.60	12,050	0.07	0.95
Gender (boys compared to girls)	-2.17	0.61	12,030	-3.54	<.0001*
Number of districts from K to 3rd	-1.26	0.38	12,090	-3.27	<.0001*
KRA Overall Score	1.41	0.03	12,120	46.73	<.0001*

*Note. * indicates statistical significance at p < .05*