Effects of Sustained Quality in PreK, Kindergarten and First Grade in DallasISD

By Annie Wright, Ph.D., Dylan Farmer, and Yusuf Kara, Ph.D.

August 2018

SMU | Annette Caldwell Simmons School of Education & Human Development

Center on Research & Evaluation

Executive Summary

A study conducted by SMU's Center on Research and Evaluation shows that DallasISD students in high quality PreKindergarten, Kindergarten and 1st grade classrooms outperform their peers in lower quality classrooms. The comprehensive approach to data-driven instructional coaching adopted by the district is positively influencing student outcomes.

What is CLASS?

DallasISD has adopted use of the CLASS system[1] which has three key components:

(1) a multi-dimensional framework for thinking about classroom quality that includes Classroom Organization, Instructional Support and Emotional Support
(2) collection of observational data by external SMU observers
(3) rapid use of observational data

What did the study find?

Quality in DallasISD. Quality in PreK, kindergarten and first grade, as measured by CLASS, is pervasive across the district. A low ratio of students are in classrooms rated as low quality and the district is actively engaged in both increasing and sustaining quality. Across all analyses, an optimal pathway is evidenced: when students are exposed to sustained quality for multiple consecutive years, their academic outcomes are best. The Emotional Support and Classroom Organization domains if the CLASS model seem to be most responsible for driving these trends.

PreK Takeaways

Enrolling in PreK of any level of quality is better than not enrolling in PreK at all and a high quality PreK environment produces the best results. Students in high quality PreK classrooms are more likely to be academically ready for kindergarten. This readiness translates into a greater likelihood of being successful during the kindergarten year, which in turn sets a strong foundation for early elementary experiences. CORE's analyses show that especially when Emotional Support is present in the PreK classroom, students benefit.

Kindergarten Takeaways

Quality kindergarten should be paired with quality PreK for optimal end-of-kinder results, especially on language arts and reading. Quality PreK followed by quality kindergarten makes students significantly more likely to be academically "on track" at the end of kindergarten. High quality kindergarten cannot compensate for lack of a high quality PreK experience and low quality kindergarten can erode some of the positive gains achieved by high quality PreK. The positive effects of high quality kindergarten *add-to* the positive effect of high quality PreK but do not make up for the absence of it.

First Grade Takeaways

PreK quality continues to have positive influence on academic outcomes through the first grade; over time, the domains of Emotional Support and Classroom Organization seem to be the drivers of these trends. The effect of high quality Emotional Support and Classroom Organization during PreK shows a consistent and positive effect on end-of-1st grade outcomes for language arts and reading, regardless of the quality of kindergarten and 1st grade. Demonstrating the accumulating impacts of quality over multiple years, positive effects of high quality K and 1st in these same domains continue to add-to the positive effect when PreK quality is high in these areas, but does not make up for the absence of it.

[1] Classroom Assessment Scoring System; https://teachstone.com/class/

How was the study conducted?

DallasISD's Early Learning Department and SMU's Center on Research and Evaluation (CORE) are engaged in a longitudinal research-practice partnership aimed at supporting quality early childhood educational experiences. CLASS observations are conducted by CORE and rapid feedback of actionable data is allowing the Early Learning Department, instructional specialists and campus leadership to provide targeted coaching to early elementary teachers (prekindergarten to second grade). This ongoing data collection is also allowing CORE to document how quality is changing semester over semester and year over year and to document the longer-term implications of sustained early childhood quality for student outcomes.

For the current analyses, CORE focused on a single cohort of students; those who entered PreK in the 2015-16 school year, kindergarten in the 2016-17 school year, and first grade in the 2017-18 school year. Data sources used to conduct the analyses include: extant DallasISD data about student enrollment, demographic characteristics and academic outcomes (Istation's ISIP literacy assessment and the Terra Nova standardized assessment), CLASS® observational data. CORE provides annual descriptive reporting to DallasISD. This is an additional and deeper-dive report, made possible by local funders, focusing on PreK, Kindergarten and first grade analyses. CORE will continue to create supplemental reports as additional longitudinal data become available and as cohorts of students move through subsequent years of elementary school.

Why does it matter?

CORE's findings support other national evidence that the PreK years represent a critical opportunity for learning that is not easily regained if missed; excellent Kindergarten cannot make up for having no or low quality PreK in terms of end-of-Kindergarten academic assessments. Other national data shows that these effects persist over time; over time, CORE will continue to assess impacts on student outcomes through grade levels.

Some national studies point to a "fade out" effect where the influence of PreK diminishes over time. However, many of these design lack sustained data about quality and therefore provide only a partial picture of how PreK really works. The DallasISD data is unique in the nation; it is a large sample that represents a diverse student body, it utilizes external and highly trained observers utilizing a psychometrically valid rating tool, and it defines quality as aspects of the transactional relationship between students and teachers and supportive aspects of the classroom climate.

The CLASS approach avoids some of the biggest pitfalls of high stakes testing and the accountability mindset that is pervasive in education, while still setting high expectations and then supporting teachers and school leaders in reaching them. CLASS does not conflate student academic performance and teacher or classroom quality. CLASS assesses classroom quality in and of itself. These conditions for quality do not depend on the characteristics of students that walk in the door and instead focus on specific and malleable actions that are within teachers' and school leaders' control. These conditions for quality can then be related to student outcomes in order to determine broad impacts for academic achievement.

These findings point to a sustainable and scaleable way forward. That is, the comprehensive adoption of the CLASS system by DallasISD is demonstrating how quality can be provided, at scale. The DallasISD Early Learning Department has adopted a comprehensive approach to instructional coaching that has created broad conceptual aims that all can orient toward, balanced with more granular drill-down instructional coaching, customized to individual teachers or school settings. In partnership with CORE, DallasISD is collecting data for ongoing formative assessment and continuous quality improvements with early learning teachers. The ongoing research partnership demonstrates a proof of concept for ongoing quality improvement, and data-informed decision making.

Background

Several key facts about prekindergarten (PreK) nationally and in DallasISD help set the stage for understanding the rationale for CORE's analyses and the implications of the findings.

National Trends in PreK

One PreK year is necessary but not sufficient for later outcomes.

Long term learning trajectories depend on the quality of learning experiences both before and after the PreK year[i,ii,iii]. Our collective understanding of the conditions under which the effects of early experiences are "accentuated, attenuated or maintained"[iv] is limited for two key reasons yet the DallasISD data addresses both limitations.

First, in many studies of PreK impacts on third grade academic outcomes, the PreK year was considered in isolation from students' experiences in kindergarten (K), first and second grade; data about the quality of experiences following the PreK year is simply not collected[v]. In these instances, the PreK year is projected to impact the distal third grade outcome without accounting for the intervening experience.

Second, in instances where data about the intervening years (K, first and second) is collected, studies have relied on conveniently available data to approximate "quality" of the learning environment[vi]. For example, extant or structural data about the type of elementary program, teachers' credentials or class size, has been used to approximate quality. These data have produced mixed results[vii,viii]: some studies show sustained impacts over time while others show fade-out of effects above and beyond what would be expected with natural duration of time.

Quality matters, and quality is all about relationships.

More recent studies have built on emerging knowledge that the nature of children's interactions with teachers are cornerstones for robust learning[ix]. These studies operationalize quality as a process and as a feature of relationships and interactions between children and adults in the learning environment.

Ansari and Pianta (2018) utilized a measure of quality that fits this description in order to monitor the effect of sustained quality on long-term outcomes for students. With a population of largely middle class and Caucasian children, they found that when quality was consistently present across elementary grades, students' outcomes were better. Additionally, when lower quality PreK is followed by higher quality elementary school experiences, student outcomes were not as strong as sustained quality throughout all years. Subsequent experiences in high quality elementary school settings had few benefits in the absence of high quality early care. Simply put: it seems that even excellent lower elementary experiences cannot compensate for missed or low quality early childhood experiences.

When we conceptualize classroom quality as more than just features of a program or curriculum, and as inextricable from interactive relationships we must also consider the desired outcomes for students more broadly and more relationally[x]. Quality refers to positive and proximal interactions between adults and children in a learning environment. These interactions form the bedrock of academic, social and emotional skills for very young children that facilitate long-term learning.

While early literacy is a critical and key outcome of PreK and early childhood programs, it is not the only desired impact. Other academic domains, including math, and social and emotional outcomes are likely equally important as early indicators of long term and positive success for students[xi]. By focusing only on literacy as an outcome of early learning, we may be, collectively, missing key positive impacts of quality PreK.

Implementation matters too.

The ultimate impacts of any initiative are fully contingent on how well it is implemented, and early childhood education is no exception[xii]. Some key supports for implementation serve a critical if indirect role in supporting later outcomes for students. Among these is a comprehensive coaching program that supports teachers and focuses on characteristics of the process not just the structure of the PreK program[xiii], is tied to a coherent underlying framework, and promotes orderly but active classrooms leads to better outcomes for students[xiv].

PreK in DallasISD

The DallasISD Early Learning Department is a microcosm of what is happening nationally in early childhood education. Recognizing the pivotal role that early learning can play for students, DallasISD has embarked on a comprehensive PreK initiative to increase both enrollment and quality. Further recognizing that a quality PreK year alone cannot address long-term outcomes for students, DallasISD is focusing on PreK for 3-year-olds and 4-year-olds through second grade as an early learning continuum. The commitment to monitor the process quality of the early learning environment using the CLASS® tool means that a) timely formative information is being gathered and shared for purposes of instructional coaching and b) many of the gaps in knowledge seen in the national literature are addressed in Dallas.

CORE's work with DallasISD presented a unique opportunity to address the broader national trends for at least two reasons: 1) the DallasISD student sample is predominantly made up of ethnic minorities and English Language Learners who are impoverished in comparison to national samples, and 2) in keeping with emerging theory and practice, quality is measured with the CLASS® tool and is operationalized as characteristics of the student-teacher relationship and classroom climate.

Methods

Research Design

CORE obtained approvals from SMU's Institutional Review Board and DallasISD's Research Review Board to conduct this research. CORE used a broad analytical strategy to explore the optimal pathways for success in the early elementary years.

Sample

In DallasISD students are eligible for full-day PreK if they are an English Language Learner (ELL), in the foster care system, part of a military family, homeless, or economically disadvantaged. Additionally, a small cohort of non-eligible students attend PreK in DallasISD through an optional payment program.

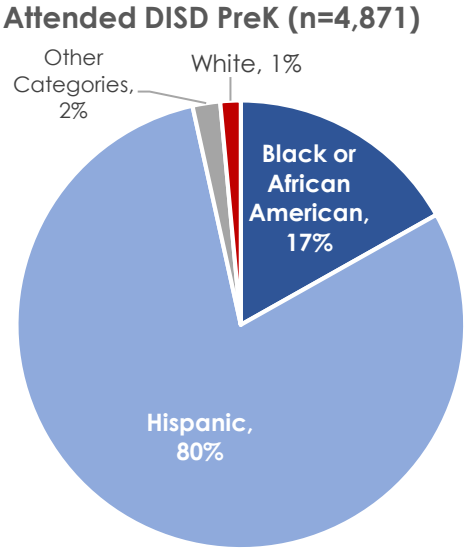
For the current analyses, CORE limited the sample based on two considerations.

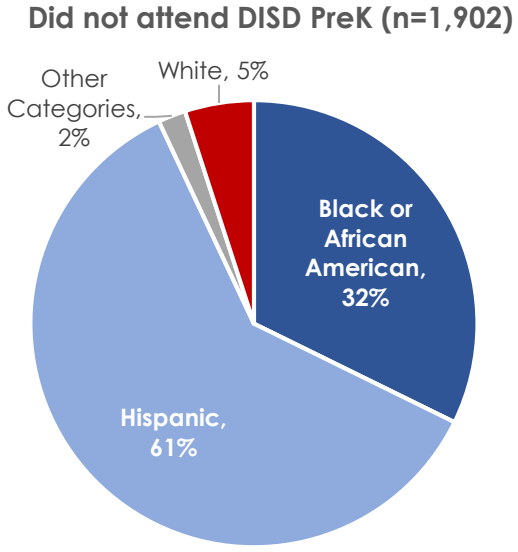
First, the sample included only economically disadvantaged students enrolled in DallasISD PreK in the 2015-16 school year, attended DallasISD K in the 2016-17 school year, and attended DallasISD first grade in the 2017-18 school year[xv]. The comparison group includes economically disadvantaged students who attended DallasISD K and first grade, but not DallasISD PreK. These three consecutive years were chosen for analyses as they were the years that DallasISD and SMU CORE started collecting CLASS® observation data in order to promote and measure quality. During these years, all PreK classrooms were observed and all K, first grade and second grade classrooms in approximately 45% (68 of the 147 elementary campuses) were observed.

Second, the sample was limited to students who had valid K readiness data. The table below describes the overall group of students who had valid K readiness data available for initial analysis of the effect of PreK. Note that this does not represent all DallasISD 2016-17 K students who did or did not attend PreK. Sample sizes applicable for the analyses described in this report vary according to availability of the CLASS® quality and student outcome data required for each analysis.

Table One. Description of DallasISD 2015-16 PreK – 1st Grade Sample used in analyses

Gender		
	Male	3,377 (50.1%)
	Female	3,396 (49.9%)
Attended DISD PreK		
	Yes	4,871 (71.9%)
	No	1,902 (28.1%)
Race		
	American Indian or Alaska Native	7 (0.1%)
	Asian	44 (0.7%)
	Black or African American	1439 (21.3%)
	Hispanic	5055 (74.6%)
	Unknown or Multiple	61 (0.9%)
	Native Hawaiian/Pacific Islander	1 (0.01%)
	White	166 (2.5%)





Measures

The data sources used to conduct the analyses were:

- **Extant DallasISD data about student enrollment and demographic characteristics.** This includes the teacher of record for each student during each school year, school attended, daily attendance, risk-factor indicators, race/ethnicity, and status as an English Language Learner. This enrollment data was used to link student outcomes to the CLASS® scores of the PreK-second grade classroom in which each student was enrolled.

- **Extant DallasISD data about academic outcomes.**
  - Istation's Indicators of Progress (ISIP) – Early Reading. This includes beginning, middle and end of year scale scores and tier placement for each student who completed the assessment at each time point. This computer-adaptive assessment is taken by students three times per year in grades K-2 and assesses early literacy skills such as alphabetic knowledge, phonemic awareness, vocabulary, and listening comprehension. Students are placed into a performance tier according to their percentile rank. Students

scoring at the 40th percentile or above (as compared to what is "normal" for their age level) are considered to be Tier 1, or "on-track".
    - TerraNova and Supera (the Spanish version of TerraNova) are normed-referenced achievement measures used to assess DallasISD students in grades K-2. Students are assessed once-per-year, at the end of the school year, in the reading/language arts and mathematics content areas. These results compare individual student achievement relative to a nationwide normative comparison group.

- **CLASS® observation data.** CORE utilizes a diverse team of certified CLASS® observers to observe all PreK and a selection of K-2 classrooms in DallasISD. These classroom-level scores rate classroom teachers on a scale of 1-7, with 7 being the best score possible, on the quality of effective teacher-child interactions that facilitate learning and development. Scores are divided into 3 domains – emotional support, classroom organization, and instructional support – and each of these three domain scores are categorized as "meeting quality" or "not meeting quality". Therefore, a single classroom may meet the quality standard in zero, one, two, or all three domains. Classrooms that meet the quality standard in all three domains are considered to be high quality and classrooms that meet the standard in none of the three domains are considered to be low quality.

Findings

Four main analyses are described in this section of the report. A summary of these findings is provided in Table Two.

Table Two. Summary of Findings

	Brief Description	Takeaway
Analysis 1	Focuses on quality at each grade level in isolation and conceptually links the year over year findings; uses quality thresholds and examines probability of being "on track" - Tier 1 on ISIP early reading	The optimal pathway and greatest likelihood of being on track and "ready" for each new grade is having a quality experience for multiple consecutive years
Analysis 2	Describes the multiplicative effect of two and three years of quality (PreK and K; PreK, K, and 1st); uses quality thresholds and examines TerraNova/Supera national percentile scores (NPC) for math, language arts and reading at the end of K and again at the end of 1st grade	Quality kindergarten should be paired with quality PreK for optimal end-of-kinder results, especially on language arts and reading; High quality K does not appear to "make up" for missing out on high quality PreK. By the end of first grade, high quality K and 1st does not appear to make-up for missing high quality PreK for mathematics outcomes.
Analysis 3	Models and tests the unique effects of PreK, K, and 1st grade quality at the end of kindergarten and 1st grade; groups quality in to high, medium and low categories and examines TerraNova/ Supera national percentile scores (NPC) for math, language arts and reading at the end of K and again at the end of 1st grade.	The effect of high quality PreK shows a consistent and positive effect on end-of-K outcomes regardless of the quality of K; positive effects of high quality K add-to the positive effect of high quality PreK but do not make up for the absence of it.
Analysis 4	Explores of the unique contribution of the different domains of quality on 1st grade outcomes.	The effect of high quality Emotional Support and Classroom Organization during PreK shows a consistent and positive effect on end-of-1st grade outcomes for language arts and reading, regardless of the quality of kindergarten and 1st grade; positive effects of high quality K and 1st in these same domains of Emotional Support and Classroom Organization add-to the positive effect when PreK quality is high in these areas, but does not make up for the absence of it.

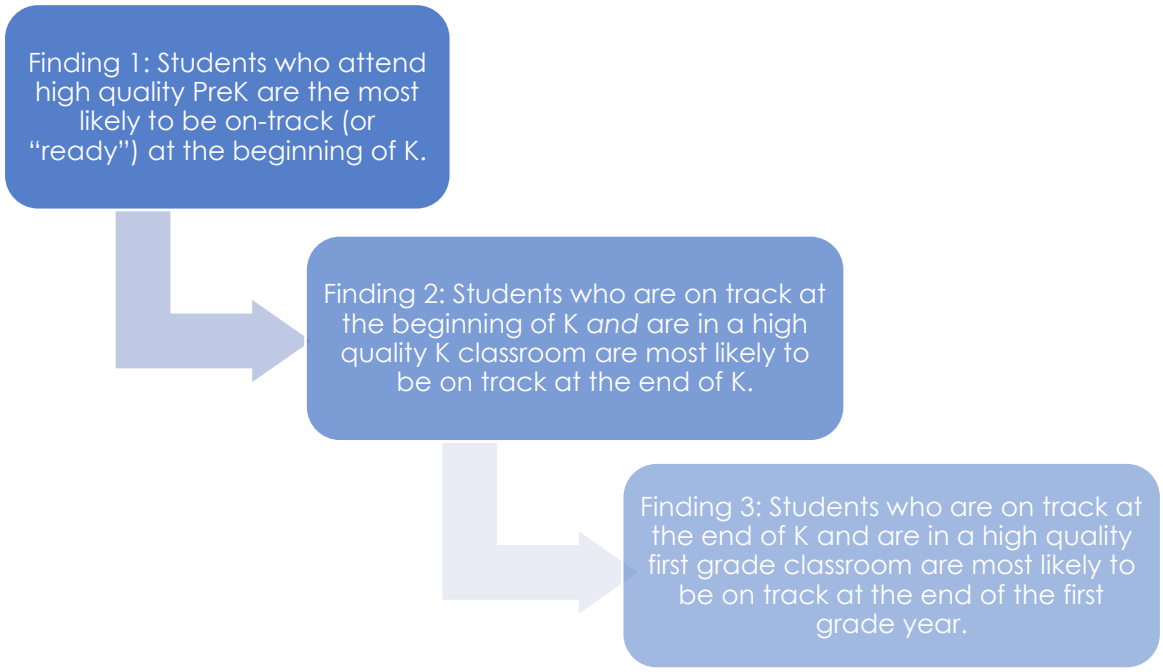
Analysis #1

For the first set of analyses, CORE asked a series of sequential questions, probing for an optimal pathway that a student might take through the early elementary years in DallasISD (see Figure One). Results for Analysis #1 focus on ISIP.

A series of logistic regression analyses were used to determine the probability of a student being "on-track" academically given their prior academic performance, their PreK experience, and their K and first experiences. In all analyses, these factors were all significant contributing factors (at $p<0.001$) to the probability that a student would be performing on-track (in Tier 1 on ISIP-ER) at each time point. Below, each finding is described in more detail and Figures One and Two and Appendix A present detailed results. This first series of analyses seek to answer the following question and sub-questions:

1. When students are exposed to higher quality PreK, are they more likely to be ready for K?
    a. When students are on-track ("ready") for K and are exposed to higher quality K, are they more likely to remain on-track by the end of K?
    b. When students are on-track at the end of K and are exposed to higher quality first grade, are they more likely to remain on-track by the end of the first grade year?
    c. How do these pathways differ for different types of students?
    d. What is the optimal pathway?

Figure One. Summary of Findings



Finding 1.1: Students who attend high quality PreK are the most likely to be on-track (or "ready") at the beginning of K. Out of a total study sample of 6,773 students, a subgroup of 3,905 students (who were in K in 2016-17) were included in this analysis. This subsample includes students who have an ISIP performance tier from the beginning of K and either did not attend PreK *or* attended PreK and have a valid CLASS® score assigned to that student's PreK classroom. Of the 3,905 students in the sample, 48.7% (n=1,902) did not attend PreK at all, 42.8% (n=1,672) were in high-quality PreK classrooms, and 8.5% (n=331) were in low-quality PreK classrooms.

The dependent variable is whether the student was in Tier 1 at the beginning of K; "yes" is equal to 1 if the student was classified as K-ready, and 0 otherwise. The results from the first model indicate that students who attend DallasISD PreK are more likely to be K-ready; attending PreK is a significant predictor of being K-ready at the 0.001 level. The odds of being K-ready are 3.3 times higher for students who attended PreK than those who did not.

Further, for students who did attend PreK, the odds of being K-ready increase 1.1 times for each of the three quality standards met by the PreK classroom the student attended.

Students who did not attend PreK were 33.1% likely to be on-track or "ready" for K, meaning they scored in Tier1 on ISIP at the beginning of K (26% for African American students and 37% for Hispanic students). Students did attend PreK but whose classrooms were rated at a low quality were 53.5% likely to be "ready" (49% for African American students and 56% for Hispanic students). Finally, students who attended a high quality DallasISD PreK classroom were the most likely to be ready for K at 61.7% likelihood (56% for African American and 63% for Hispanic).

Finding 1.2: Students who are on-track at the beginning of K *and* are in a high quality K classroom are most likely to be on-track at the end of K. Next, CORE examined whether students who were ready for K were more likely to be on-track at the end of K and whether K classroom quality also increased this likelihood for both K-ready non-K-ready students. First, we examined the extent to which being K-ready was associated with being on-track at the end of the K year. A total of 6,045 students were included in this analysis. Results from the first model indicate that students who are K-ready have a 5.5 times higher odds of being on-track at the end of K compared to those who were not K-ready; K readiness significantly predicts being on-grade level at the end of K at the 0.001 level.

To further determine whether quality of the K classroom was associated with being on-track at the end of K, a second and third model tested the degree to which the number of quality indicators met by the K classroom predicted being "on-track" at the end of the year for both the K-ready and non-K-ready students. For both groups, the odds of being K-ready increased significantly for each of the three quality standards met by the K classroom. For students who were not K-ready, the odds of being on-track at the end of K increased 1.5 times for each quality standard met by the K classroom. The odds increased 1.4 times for K-ready students. This illustrates that K quality does not have an exponential effect on students who are not on-track on the beginning of K. In other words, their odds of "catching up" by the end of K increase as K quality increases, but not at a rate greater than that of their K-ready peers.

For the 2,673 students in this study sample who were not on-track at the beginning of K, 4.4% (n=117) were in a classroom with a low quality CLASS® rating and 9.4% (n=251) were in a high quality classroom. The rest were in medium quality classrooms (i.e., those that met 1 or 2 quality standards out of the possible 3). Non-K-ready students in low quality K classrooms were just 19.7% likely to be on-track at the end of the K year (14% for African American and 27% for Hispanic). Conversely, non-K-ready students in high quality K classrooms were 49.4% likely to be on-track at the end of K.

For the 3,185 students in this sample who were K-ready, just 2.2% (n=69) were in a low quality K classroom. These students, who *were all on-track at the beginning of K*, were 58% likely to remain on-track by the end of K (46% for African American students and 66% for Hispanic). Of the remaining K-ready students, 10.9% (n=349) were in high quality K classrooms (the remaining sample of K-ready students were in medium quality classes). These K-ready students who had a high quality K classroom had the highest likelihood of being on-track at the end of K at 82.8% (76% for African American students and 84% for Hispanic).

Finding 1.3: Students who are on-track at the end of K *and* are in a high quality first grade classroom are most likely to be on-track at the end of the first grade year. Finally, CORE examined whether students who were on-track at the end of K were more likely to be on-track at the end of first grade and whether first grade classroom quality also increased this likelihood for both on-track and not-on-track students. First, we examined the extent to which being on-track at the end of K was associated with being on-track in the end of the first grade year. A total of 5,279 students were included in this analysis. Results from the first model indicate that students who were on-track at the end of K have 6.9 times higher odds of being on-track in the end of first grade compared to their peers that were not on-track by the end of K; end-of-K performance significantly predicts being on-grade level in the end of first grade at the 0.001 level.

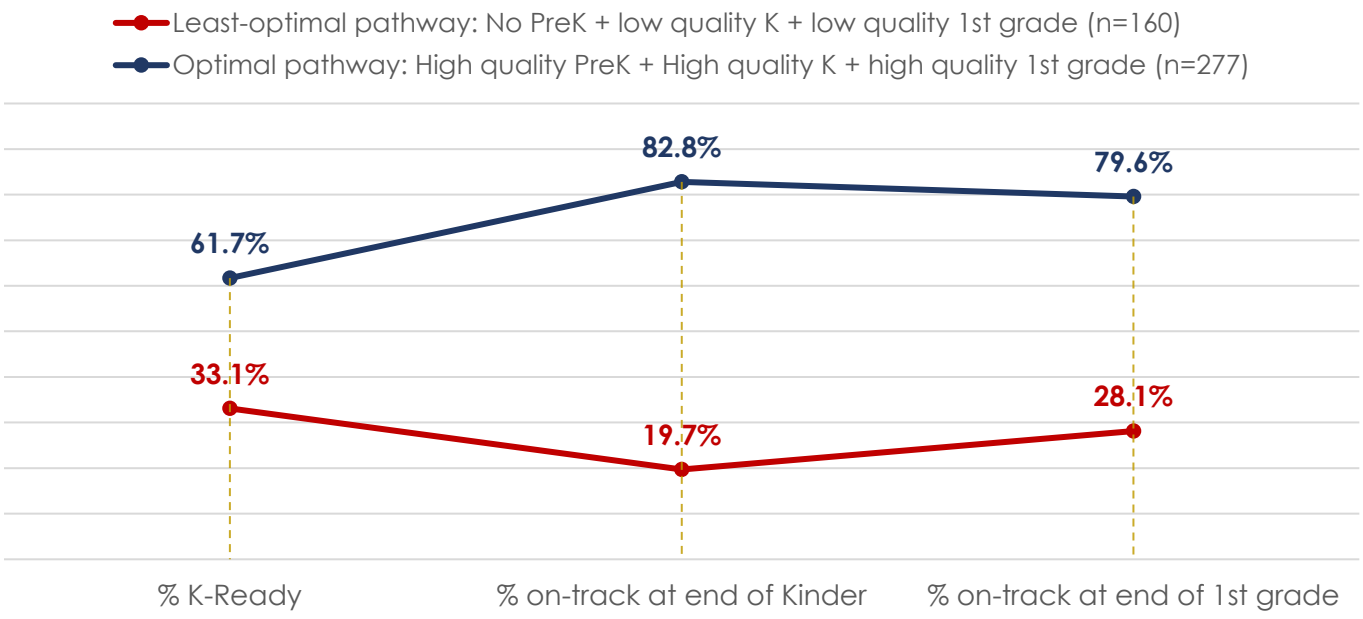
To further determine whether quality of the first grade classroom was associated with being on-track at the end of first grade, a second and third model tested the degree to which the number of quality indicators met by the first grade classroom predicted being "on-track" at the end of the year for both the previously on-track and not-

on-track groups of students. For both groups, the odds of being on-track at the end of first grade increased significantly for each of the three quality standards met by the first grade classroom. For students who were not on-track at the end of K, the odds of being on-track at the end of first grade increased 1.12 times for each quality standard met by the first grade classroom. The same odds increased 1.18 times for students who were on-track at the end of K. This finding once again illustrates that early elementary quality does not have an exponential effect on students who are not on-track on the beginning of any given school year. In other words, their odds of "catching up" by the end of K or the end of first grade increase as classroom quality increases, but not at a rate significantly greater than that of their on-track peers.

For the 2,591 students in this study sample who were not on-track at the end of K, 12.9% (n=334) were in a classroom with a low quality CLASS® rating and 20.5% (n=532) were in a high quality classroom. The rest were in medium quality classrooms (i.e., those that met 1 or 2 quality standards out of the possible 3). Students who were not on-track at the end of K and then enrolled in low quality first grade classrooms were just 28.1% likely to be on-track by the end of the first grade year (25% for African American and 32% for Hispanic). Conversely, students who were not on-track at the end of K but then enrolled in a high quality first grade classrooms were 35.3% likely to be on-track by the end of first grade.

For the 3,206 students in this sample who were on-track at the end of K, just 9.4% (n=302) were in a low quality first grade classroom. These students, who *were all on-track at the end of K*, were 69.5% likely to remain on-track by the end of first grade (63% for African American students and 74% for Hispanic). Of the remaining students on-track at the end of K, 23.5% (n=755) were in high quality first grade classrooms (the remaining sample of on-track students were in medium quality first grade classes). These students who were on-track at the end of K and then were in a high quality first grade classroom had the highest likelihood of being on-track by the end of first grade at 79.6% (66% for African American students and 83% for Hispanic).

Figure Two: % Likelihood of Starting & Staying On-Track through the Early Grades in Most-Optimal and Least-Optimal Pathways



Key Takeaways from Analysis #1

Two key takeaways are important to highlight and interpret further.

- First, even if students enter kindergarten ready, a low quality kindergarten experience can lessen the previous and positive effects of quality PreK. Students who entered kindergarten "ready" (they had scored in Tier 1 on ISIP at the beginning of K) but who were in low quality K classrooms significantly under-performed peers who had also been ready at the beginning of K but who were in high quality K classrooms. Just 58% of those who entered K on-track but were in a low quality K classroom were still on-track by the end of the year, compared to 82.8% of those who also entered on-track and were in a high quality classroom. This is some of the strongest available local data to date to suggest that a lower quality K classroom can plausibly erode some of the positive gains that were realized due to a high quality PreK year and points to the need for sustained quality across consecutive years. A similar trend was seen in the first grade year but the magnitude of the difference was not as strong. Additionally, these findings suggest that high quality K and first grade are not enough sufficient alone to "catch-up" students who do not enter K "ready", and attending high quality PreK gives students significantly higher odds of achieving K readiness.

- Second, the series of answers that emerge from these analyses point to an optimal pathway for students and reinforce the critical role that sustained quality of multiple years of quality instruction can have. In Figure Two, two different pathways that students have taken through early elementary in DallasISD are highlighted. In the best or optimal pathway, students accumulate advantage by attending a high quality PreK classroom, are "on-track" or ready at the beginning of their K year, attend a high quality K classroom, are on-track at the end of K and attend a high quality first grade classroom. On this optimal pathway, 62.4% of students are on-track at the beginning of K, 82.7% are on-track at the end of K, and 79.6% are on-track at the end of the first grade year. In the least optimal pathway, students do not attend PreK at all, are not on-track or ready at the beginning of K, attend a lower quality K classroom, are not on-track at the end of K, and attend a lower quality first grade classroom. On this pathway, 33.1% of students are on-track at the beginning of K, 19.7% are on-track at the end of K, and just 28.1% are on-track by the end of first grade.

Analysis #2

The second series of analyses probed further in to the trends established in Analysis 1. This analysis used distributional comparisons to answer the questions:

- Does sustained quality over consecutive years (PreK, K, and 1st grade) lead to the best outcomes for students?
- Is high quality PreK alone enough to achieve these outcomes?
- Can high quality K-2 "catch-up" students if they didn't receive high quality PreK?

First, Analysis #2 findings are summarized, then the results and a further description of the analyses are provided. Visual representations of the results of these analyses using the ISIP English and Espanol scale scores at the end of K and 1st grade are provided in Appendix B.

Although these comparisons give us good impressions about the student outcome-classroom quality relationships, complex models are needed for a deeper evaluation, and of course for statistical testing. Thus, in analysis #3, we will present the results of a series of linear models that are analyzed to interpret the relationships between the classroom quality and student achievement.

Key Takeaways from Analysis #2 for Kindergarten

Finding 2.1: Students who receive both high quality PreK and high quality K perform better on language arts and reading at the end of K as compared to their peers that did not go to PreK and had low quality K; on math scores, high quality PreK and lower quality K produced the best outcomes. These results are strongest for language arts and reading. Although comparisons up this point imply that the group who had combined high quality PreK and

K had higher means, medians, and % of students with 40+ NPC, we cannot examine the unique effect of PreK and/or K quality on the student outcomes, based on presented graphs. In other words, these analyses examine the worst and best-case scenarios for the classroom quality, without focusing on the unique contributions of PreK and K. Results in the following section can give us some impressions about the unique effects of PreK and K quality.

Finding 2.2: High quality K does not appear to "make up" for missing out on high quality PreK. All three-group distributional comparisons showed that groups who had at least one high quality education experience of either PreK or K, as defined by CLASS®, had higher outcome indicators (mean, median, and % of 40+ NPC) compared to the reference group, who did not attend DallasISD PreK and had low quality K. Moreover, distributional comparisons also showed that the group who had low quality K, but had high quality PreK, had slightly higher indicators of success, compared to the group who had high quality K but did not go to DallasISD PreK. Although these comparisons give us good impressions about the student outcome-classroom quality relationships, complex models are needed for a deeper evaluation, and of course for statistical testing. Thus, below we present the results of a series of linear models that are analyzed to interpret the relationships between the classroom quality and student achievement.

Analysis #2 differs from Analysis #1 in that they considered a combined or cumulative effect of multiple years, whereas the first series of analyses considered each year in isolation but in consideration of how the previous year prepared students for success. To get preliminary impressions of the multiplicative effect of multiple years of quality combined, we compared outcomes of the students how had high quality PreK and high quality K (reference group; n=242) to students who did not attend DallasISD PreK and were in a low-quality K classroom (comparison group; n=656).

Key Takeaways from Analysis #2 for First Grade

Finding 2.3: Students who receive high quality pre-K followed by high quality K and 1st grade perform better at the end of first grade compared to their peers who do not attend pre-K and receive low quality in either K, 1st or both. As with the end of K results, these end of 1st grade results are strongest for language arts and reading. It is critical to note that these specific results cannot examine the unique effect of pre-K and/or K and 1st grade quality on the student outcomes, based on presented graphs. In other words, these analyses examine the worst and best-case scenarios for the classroom quality, without focusing on the unique contributions of pre-K, K and 1st grade. Results of Analysis #3 can give us some impressions about the unique effects of pre-K, K and 1st grade quality.

Finding 2.4: When comparing the isolated effects of high quality pre-K and high quality K and 1st grade, findings suggest that students who receive high quality DallasISD pre-K outperform students who did not attend Dallas ISD pre-K regardless of the quality of the K and 1st grade classroom in some cases. For mathematics, students who attending high quality pre-K followed by low quality K and 1st grade outperformed peers that did not attend any DallasISD pre-K followed by high quality K and 1st. In other words, high quality K and 1st does not appear to "make up" for missing out on high quality pre-K for mathematics outcomes. For language arts, the opposite finding was observed. Students who received high quality K and 1st grade outperform students who receive high quality pre-K followed by low quality K and 1st grade. No differences were observed for reading.

Further explanation and histograms illustrating these findings are shown in Appendix B.1.

Analysis #3

In these series of analyses, we first modeled the unique effects of PreK and K quality on students' end-of-K achievement outcomes. Second, we modeled the unique effects of PreK, K and 1st grade quality on students' end-of-1st grade achievement outcomes. For linear model analyses, we modified the thresholds for the overall classroom quality indicators. In other words, instead of the former classification of "low" and "high" quality, we created a three-category metric of overall quality such that the three categories are "low", "medium" and "high" quality. The rule for the new overall quality categorization is as follows:

- Low quality (LQ): None of the domain qualities were satisfied
- Medium quality (MQ): One or two of the domain qualities were satisfied
- High quality (HQ): Three of the domain qualities were satisfied

It is important to note that, for PreK-based quality results, we also included the group of students who did not attend DallasISD PreK. This group was the reference group for the PreK based comparisons and is denoted as "No PK" in the figures. For this specific series of analyses, **we will control the K and 1st grade quality experienced by students and examine whether PreK quality effects still exist at the end of K and 1st grade absent of the K and 1st grade influence.** Similar to that, we also examine the unique effect of K and 1st grade quality, by ruling out the PreK quality experience. **Thus, results are interpreted as the unique effects of each grade level by ruling out the other effect.**

Following a short description of the overall findings, we describe the analyses and specific results, starting with simpler models and going to more complex models (preliminary and less-complex models are described in Appendix C). Results are organized based on the research questions, similar to previous section. In order to prevent the effects of the non-normality on the analyses, we also performed median-based quantile regression analyses for each research question.

Finding 3.1: The effect of high quality PreK shows a consistent and positive effect on end-of-K outcomes regardless of the quality of K; positive effects of high quality K *add-to* the positive effect of high quality PreK but do not make up for the absence of it. The increase of PreK quality shows a consistently increasing positive effect on outcomes from low to high quality PreK, except in the case of the reading scores. For reading, the means and medians are slightly dropped for the high quality PreK group, compared to medium quality. Despite the fact that that decrease is not dramatic, it should be examined more elaborately by looking at the domain-specific PreK quality thresholds. As will be seen in the next section, such a decrease for reading was not observed for the K quality trends. In addition, the language specific (English/Spanish) trends can also be examined to get more detailed information.

Finding 3.2: When the effects of K and 1st are not considered, a significant positive main effect for medium and high quality PreK is observed for both language arts and reading end-of-1st grade outcomes. When effects of K and 1st are controlled for, the low quality PreK group's language arts and reading scores were higher than the no PreK group, however, higher quality of PreK did not consistently elevate that effect (low and high quality PreK groups had nearly equal mean and medians). All mean and median based effect calculations were significant for medium and high level PK, K, and 1st grade qualities, with the only exception for median of high quality PK for language arts. However, the low and high level PreK quality effects did not differ in terms of when compared to the reference (no PreK) group.

Upon finding that there was no evidence that higher quality of PreK did not consistently elevate the effect of PreK, a fourth analysis was conducted to disaggregate the quality variable. Previous exploratory analyses revealed that individual domains had notably different effects on end-of-kindergarten and end-of-first grade student outcomes, implying that high quality within each domain is affecting student outcomes differently and that taking them as a combined single quality metric may be resulting in lost information and diluting findings. For a description of domain-specific quality effects, see findings for Analysis #4.

Analysis 3.1 Unique effect of PreK at end of K student outcomes

This first set of analyses seeks to answer the question: *What is the effect of PreK quality on end of K student achievement, after controlling the effect of K quality?*

In preliminary analyses, we found that students who received medium or high quality PreK had higher mean and median national percentile scores on all three subjects, compared to students who did not attend PreK (see Appendix B). For this specific question, however, we will control the K quality experienced by students and examine whether PreK quality effects still exist.

Results are summarized in Table Three and Figures Nine-Eleven. For each specific group, significance tests should be examined compared to its reference group, by controlling the quality effect of the other grade level. For example, when we examine the significance of high quality PreK, the comparison is made to the PreK quality reference group, which is the "No PreK" group, by controlling the K quality experience. Thus, for that specific example, we are comparing the mean/median end-of-K achievement of the group who had high quality PreK to the group who did not go to PreK, by ruling out their K quality experiences. Since that coefficient is significant (highlighted in the table below), we can make an interpretation such that, the mean/median of the group who had high quality PreK is significantly higher than the group who did not go to PreK in all three assessment areas, and that this difference is uniquely attributable to PreK and is not influenced by the students' K experience.

Table Three. Mean and Median based regression results for the effect of controlled PreK and K qualities

		Math		LangArts		Reading	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	0.73	> .05	2.15	< .05	1.68	> .05
	Med	4.55	< .01	6.73	< .01	5.96	< .01
	High	3.62	< .01	6.01	< .01	4.31	< .01
	K Qual. (Ref. is Low)						
	Med	2.98	< .01	4.66	< .01	4.93	< .01
	High	3.04	< .01	6.05	< .01	5.47	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	0.91	> .05	1.32	> .05	0.64	> .05
	Med	3.78	< .01	4.92	< .01	6.37	< .01
	High	4.09	< .01	5.39	< .01	3.59	< .01
	K Qual. (Ref. is Low)						
	Med	1.78	> .05	3.63	< .01	4.71	< .01
	High	2.27	< .01	4.65	< .01	5.49	< .01

The mean and median trends for the PreK quality when controlling for K quality are given below separately for each of the subjects assessed at the end of K. For each pair of figures, two separate lines represents the trend of the PreK quality effect for the groups who had low and high quality K experiences.

Figure Nine. Trend of math means and medians for PreK quality categories, by controlling K quality

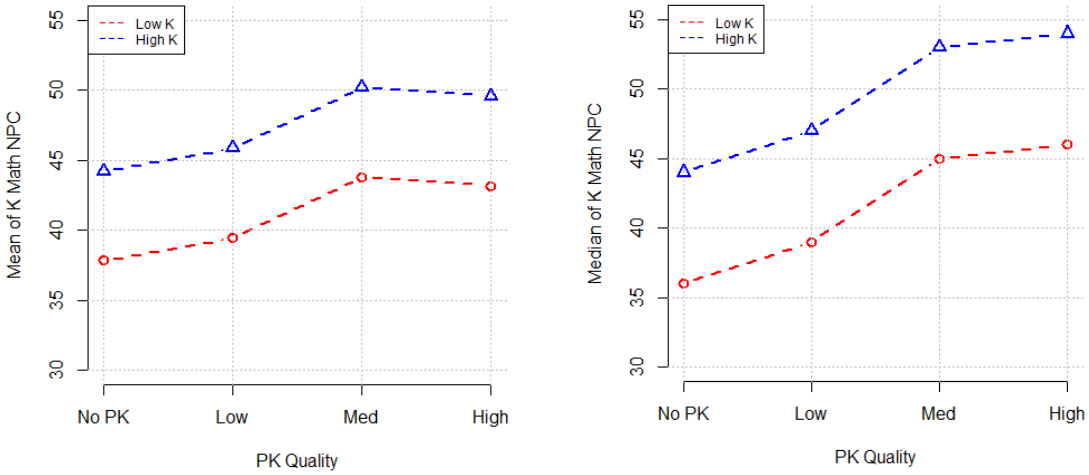


Figure Ten. Trend of language arts means and medians for PreK quality categories, by controlling K quality

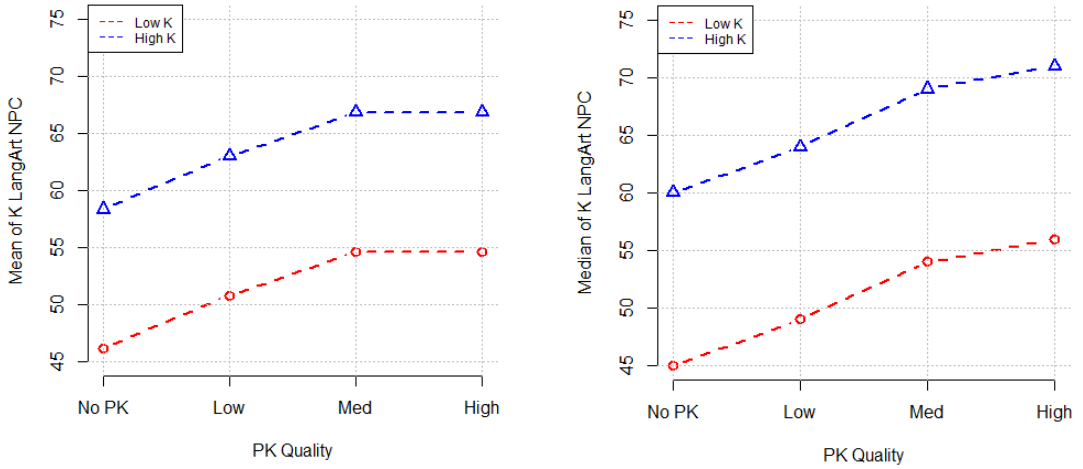
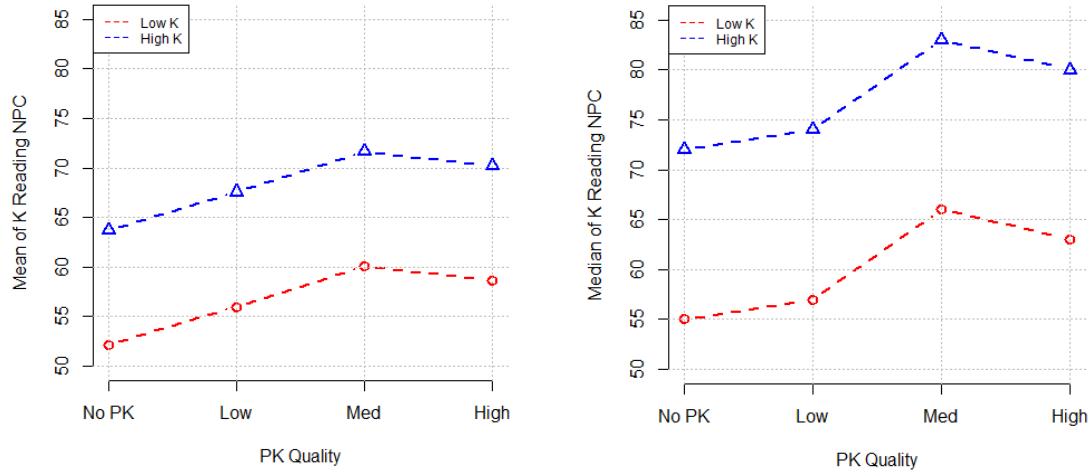


Figure 11. Trend of reading means and medians for PreK quality categories, by controlling K quality



Analysis 3.2 – Unique effect of K quality at the end of K student outcomes

This second set of analyses seeks to answer the question: *What is the effect of K quality on K student achievement, after controlling the effect of PreK quality?*

Statistical results for this question can also be followed from Table Three. We also provide similar trend graphs for K quality, namely the change of means/medians with the increase of K quality for the groups who had low and high quality PreK, respectively.

Looking at the figures 12-14, we can see that there is a consistent increase in means and medians for both the low and high quality PreK groups with the increase of K quality. Thus, we can say that the quality of K shows consistent effects on the student outcomes when coupled with both low and high quality PreK, and high quality PreK adds-to the effect of high quality K but one does not make-up for one or the other. The mean and median trends for the PreK quality when controlling for K quality are given below separately for each of the subjects assessed at the end of K. For each pair of figures, two separate lines represents the trend of the K quality effect for the groups who had low and high quality PreK experiences.

Figure 12. Trend of math means and medians for K quality categories, by controlling PreK quality

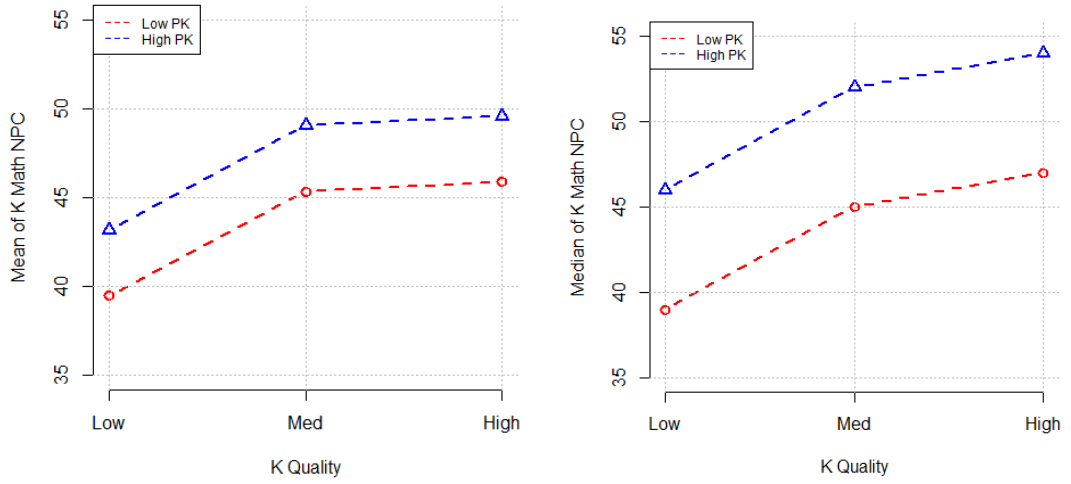


Figure 13. Trend of language arts means and medians for K quality categories, by controlling PreK quality

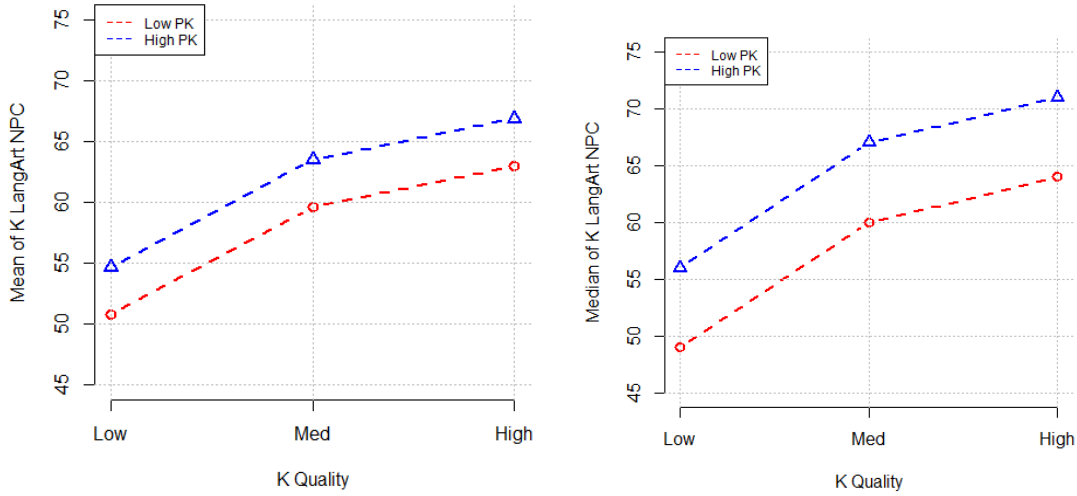
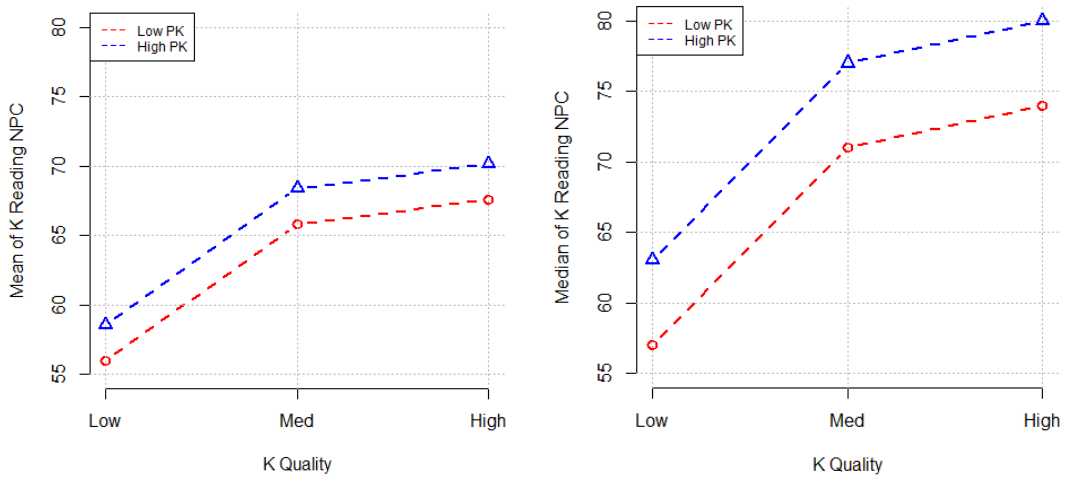


Figure 14. Trend of reading means and medians for K quality categories, by controlling PreK quality



Analysis 3.3 – Unique effect of PreK quality at the end of 1st grade student outcomes

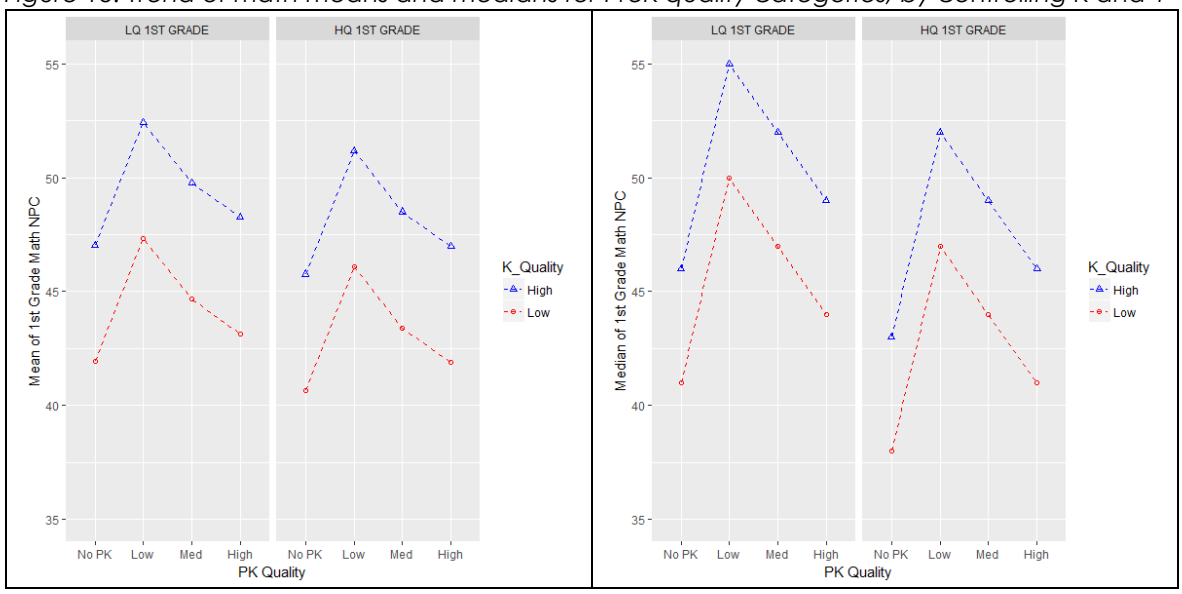
In these analyses, the partial effects of PreK, K, and 1st Grade classroom quality on end-of-1st grade math, language arts and reading NPC's were investigated. The partial effect implies that the effect of quality at each grade level was examined by controlling the effects of quality of other grade levels. Mean and median based linear model analysis results are presented in the Table Four below. This analysis is ideal, as it considers all experiences prior to the end of 1st grade rather than just one year in isolation, while providing the unique contribution of each grade level. The sample size for this linear analysis is n=2,289 first grade students.

Table Four. Mean and Median based regression results for the effect of controlled PK, K, and 1st Grade qualities

		Math		**LangArts**		**Reading**	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	2.36	< .05	1.66	> .05	2.56	< .05
	Med	1.96	> .05	3.47	< .01	4.46	< .01
	High	.78	> .05	1.97	< .05	3.70	< .01
	K Qual. (Ref. is Low)						
	Med	.70	> .05	2.36	< .05	3.63	< .01
	High	2.19	< .05	4.21	< .01	5.52	< .01
	1st Grade Qual. (Ref. is Low)						
	Med	-.24	> .05	2.44	< .05	2.62	< .01
	High	-.72	> .05	5.33	< .01	4.02	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	2.34	< .05	1.60	> .05	2.49	< .01
	Med	2.68	< .01	3.73	< .01	4.87	< .05
	High	1.29	> .05	1.71	> .05	3.21	< .01
	K Qual. (Ref. is Low)						
	Med	0.00	> .05	2.40	< .05	2.94	< .01
	High	1.22	> .05	5.35	< .01	4.97	< .01
	1st Grade Qual. (Ref. is Low)						
	Med	-.66	> .05	3.36	< .01	2.11	< .05
	High	-1.02	> .05	4.70	< .01	3.12	< .01

Regarding the PreK quality effect on 1st Grade math, the results showed that only the low quality PreK group had a significantly higher mean than the group who did not go to PreK. The medium and high quality groups, however, did not significantly differed than the no PreK group, which was an interesting finding. Median-based analyses also showed that inconsistent trend. Thus, using this "rolled-up" definition of quality, it is hard to interpret a meaningful effect of PreK quality on 1st Grade math scores, by controlling the K and 1st grade qualities. The trend for PreK quality effect on 1st Grade math scores can be observed below in Figure 15.

Figure 15. Trend of math means and medians for PreK quality categories, by controlling K and 1st Grade quality



Regarding Language Arts and Reading, the controlled PreK, K , and, 1st grade qualities showed meaningful effects on the language arts and reading scores at the end of 1st grade. All mean and median based effect calculations were significant for medium and high level PK, K, and 1st grade qualities, with the only exception for median of high quality PK for language arts. However, as can be observed in the graphs below, the low and high level PreK quality effects did not differ in terms of when compared to the reference (no PreK) group. In other words, the low quality PreK group's language arts scores were higher than the no PreK group, however, higher quality of PreK did not consistently elevate that effect (low and high quality PreK groups had nearly equal mean and medians). This fact can be observed from the non-linear trend lines from no PreK to high quality PreK in the graphs below.

Figure 16. Trend of language arts means and medians for pre-K quality categories, by controlling K and 1st Grade quality

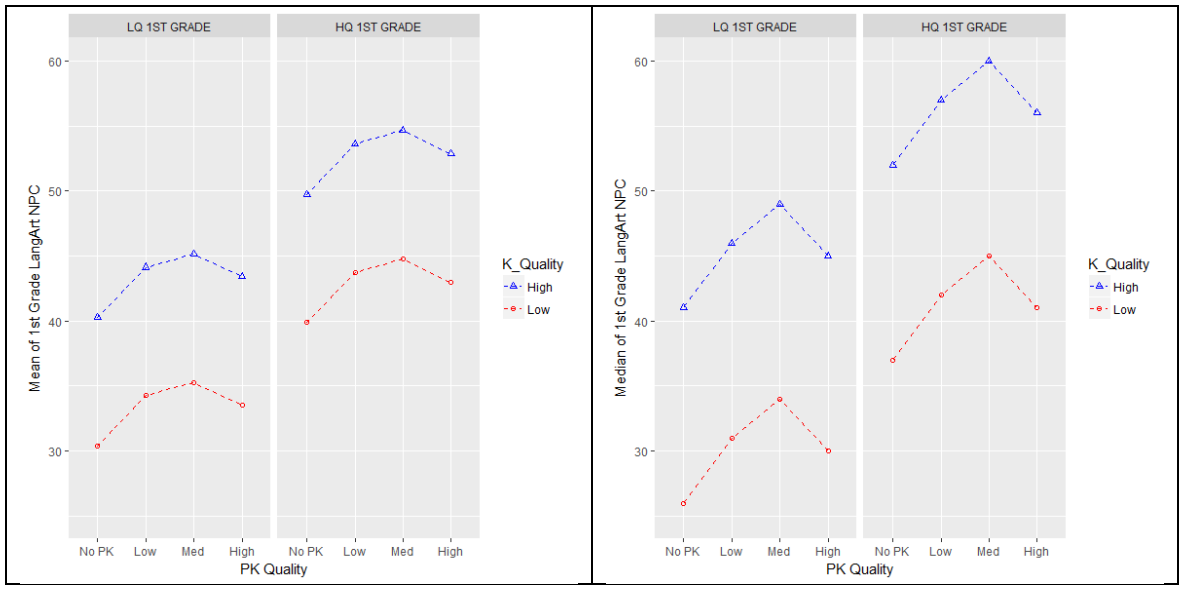
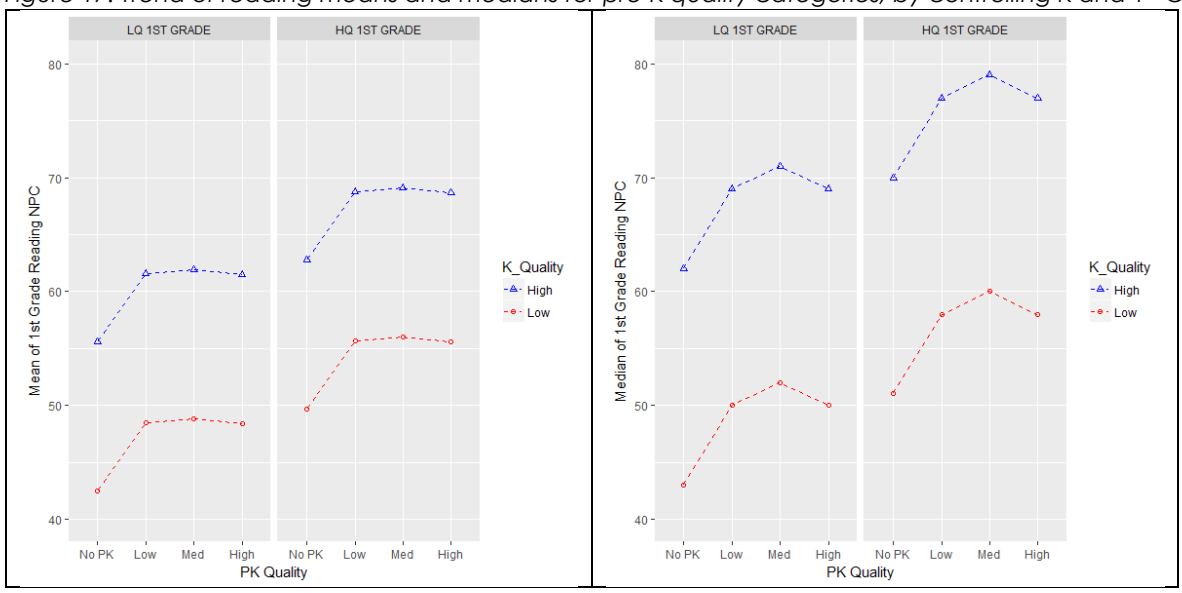


Figure 17. Trend of reading means and medians for pre-K quality categories, by controlling K and 1st Grade quality



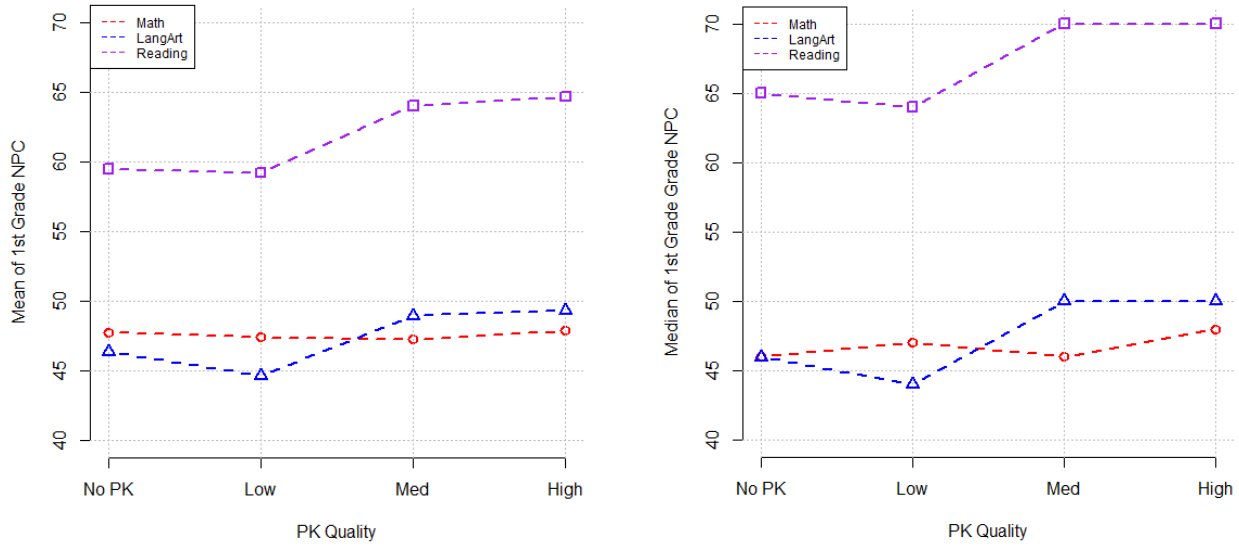
Main Effects of PK Quality at the End of 1st Grade

This series of linear models tests for main effects of PreK quality on end-of-1st grade student outcomes without controlling for the unique effect of K and 1st grade. When the effects of K and 1st are not considered, a main effect for medium and high quality PreK is observed for both language arts and reading end-of-1st grade outcomes. In other words, students who attended medium or high quality PreK outperform students who attend no PreK or low quality PreK by the end of 1st grade. No effect was observed for math.

Table Five. Mean and median based regression results for the effect of PK quality on 1st Grade NPC's

		Math		LangArts		Reading	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	-.28	> .05	-1.23	> .05	-.21	> .05
	Med	-.71	> .05	3.59	< .01	6.28	< .01
	High	.12	> .05	3.69	< .01	6.50	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	.39	> .05	-.94	> .05	-.53	> .05
	Med	0.00	> .05	3.43	< .01	4.29	< .01
	High	1.57	> .05	3.13	< .01	4.44	< .01

Figure 18. Trend of means and medians for PreK quality thresholds



Analysis #4

Upon finding that there was no evidence that higher quality of PreK did not consistently elevate the effect of PreK at the end of first grade when controlling for the contributing effect of kindergarten and 1st grade quality, a fourth analysis was conducted to disaggregate the quality variable into domain-specific effects. Previous exploratory analyses not described in this report suggested that individual domains had notably different effects on end-of-kindergarten and end-of-first grade student outcomes, implying that high quality within each domain is affecting student outcomes differently and that taking them as a combined single quality metric may be resulting in lost information and diluting findings.

In these analyses, the partial effects of PreK, K, and 1st Grade classroom quality on end-of-1st grade math, language arts and reading NPC's were investigated for each individual domain of CLASS: Emotional Support, Classroom Organization, and Instructional Support. The partial effect implies that the effect of each domain of quality at each grade level was examined by controlling the effects of that same domain of quality of other grade levels. Mean and median based linear model analysis results are presented in the Table Six below. This analysis is ideal, as it considers all experiences prior to the end of 1st grade rather than just one year in isolation, while providing the unique contribution of each grade level and the unique contribution of each domain of quality. The sample size for this linear analysis is n=2,289 first grade students.

Finding 4.1: The effect of high quality Emotional Support and Classroom Organization during PreK shows a consistent and positive effect on end-of-1st grade outcomes for language arts and reading, regardless of the quality of kindergarten and 1st grade; positive effects of high quality K and 1st in these same domains of Emotional Support and Classroom Organization add-to the positive effect when PreK quality is high in these areas, but does not make up for the absence of it. At the end of first grade, students in PreK classrooms that are high quality in Emotional Support and Classroom Organization significantly outperform their peers who attended PreK that was low quality in these domains or did not attend Pre Kata all. This is especially true for Language arts and reading and is also true for Emotional Support's effects on mathematics outcomes at the end of 1st grade.

Table Six. Mean and Median based regression results for the effect of controlled PK, K, and 1st Grade domain-specific qualities

EMOTIONAL SUPPORT		Math		LangArts		Reading	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	1.13	> .05	.48	> .05	1.56	> .05
	High	2.07	< .05	3.86	< .01	5.31	< .01
	K Qual. (Ref. is Low)						
	High	2.24	< .05	2.95	< .01	3.95	< .01
	1st Grade Qual. (Ref. is Low)						
	High	-.37	> .05	4.29	< .01	3.28	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	.26	> .05	.70	> .05	1.58	> .05
	High	1.98	< .05	3.78	< .01	5.21	< .01
	K Qual. (Ref. is Low)						
	High	1.59	> .05	2.53	< .05	2.14	< .05
	1st Grade Qual. (Ref. is Low)						
	High	-.84	> .05	3.37	< .01	2.95	< .01

INSTRUCTIONAL SUPPORT		Math		LangArts		Reading	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	2.48	< .05	3.77	< .01	4.87	< .01
	High	.67	> .05	2.03	< .05	3.80	< .01
	K Qual. (Ref. is Low)						
	High	2.72	< .05	4.18	< .01	4.62	< .01
	1st Grade Qual. (Ref. is Low)						
	High	-.72	> .05	5.01	< .01	2.93	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	2.73	< .05	3.72	< .01	4.62	< .01
	High	1.26	> .05	1.61	> .05	3.34	< .01
	K Qual. (Ref. is Low)						
	High	2.09	< .05	5.55	< .05	4.92	< .01
	1st Grade Qual. (Ref. is Low)						
	High	-1.09	> .05	3.60	< .01	2.85	< .01

CLASSROOM ORGANIZATION		Math		LangArts		Reading	
Mean Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	2.05	< .05	1.06	> .05	1.59	> .05
	High	1.88	> .05	4.00	< .01	5.65	< .01
	K Qual. (Ref. is Low)						
	High	2.27	< .05	1.55	> .05	1.98	< .05
	1st Grade Qual. (Ref. is Low)						
	High	-.31	> .05	5.28	< .01	4.77	< .01
Median Based	PK Qual. (Ref. is No PK)	t	p	t	p	t	p
	Low	1.61	> .05	.48	> .05	2.31	< .05
	High	1.95	> .05	3.58	< .01	4.79	< .01
	K Qual. (Ref. is Low)						
	High	2.26	< .05	1.04	> .05	1.90	> .05
	1st Grade Qual. (Ref. is Low)						
	High	-.44	> .05	5.27	< .01	4.19	< .01

Figure 19. Trend of math means and medians for pre-K quality domains, by controlling K and 1st Grade quality

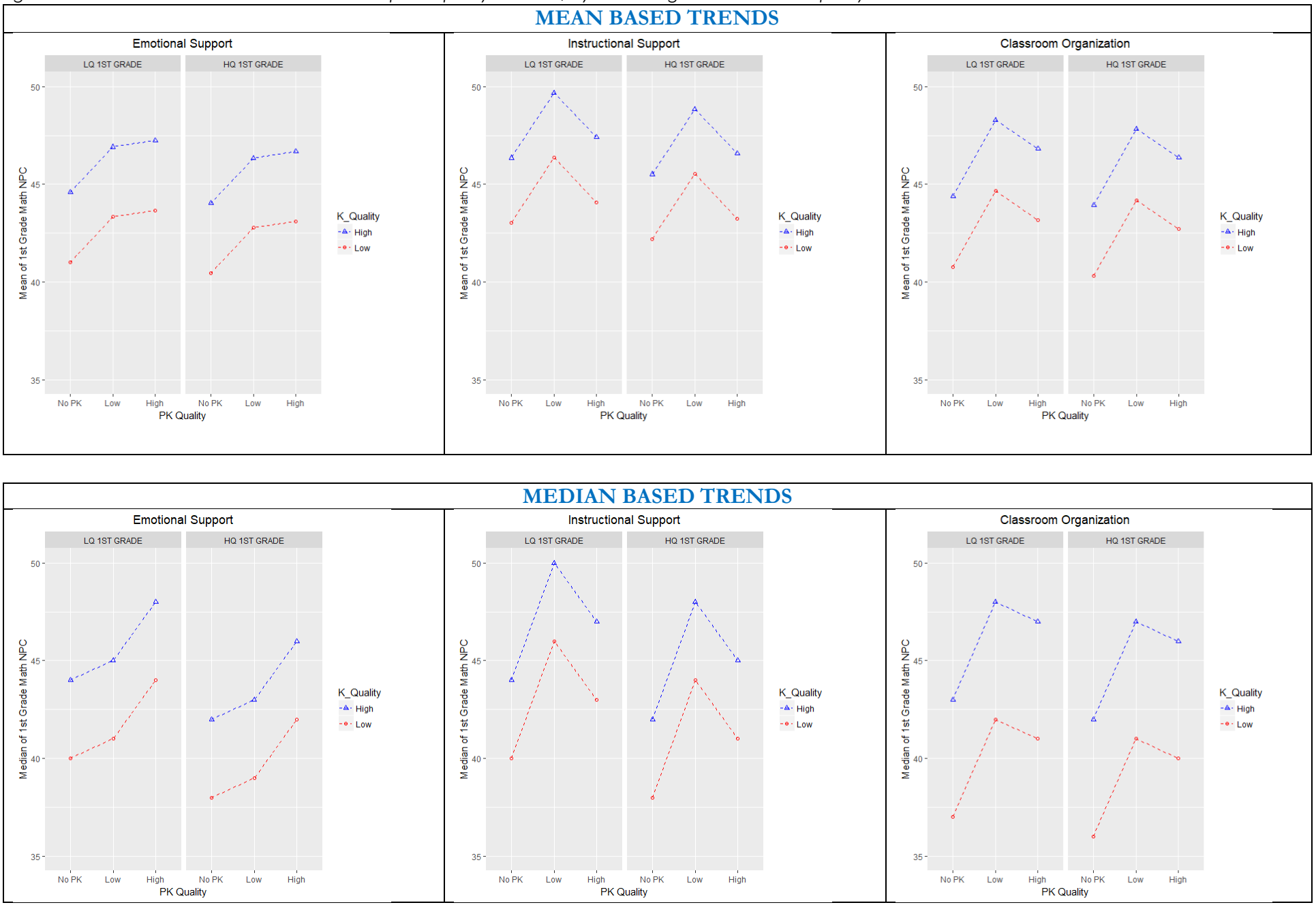


Figure 20. Trend of language arts means and medians for pre-K quality domains, by controlling K and 1st Grade quality

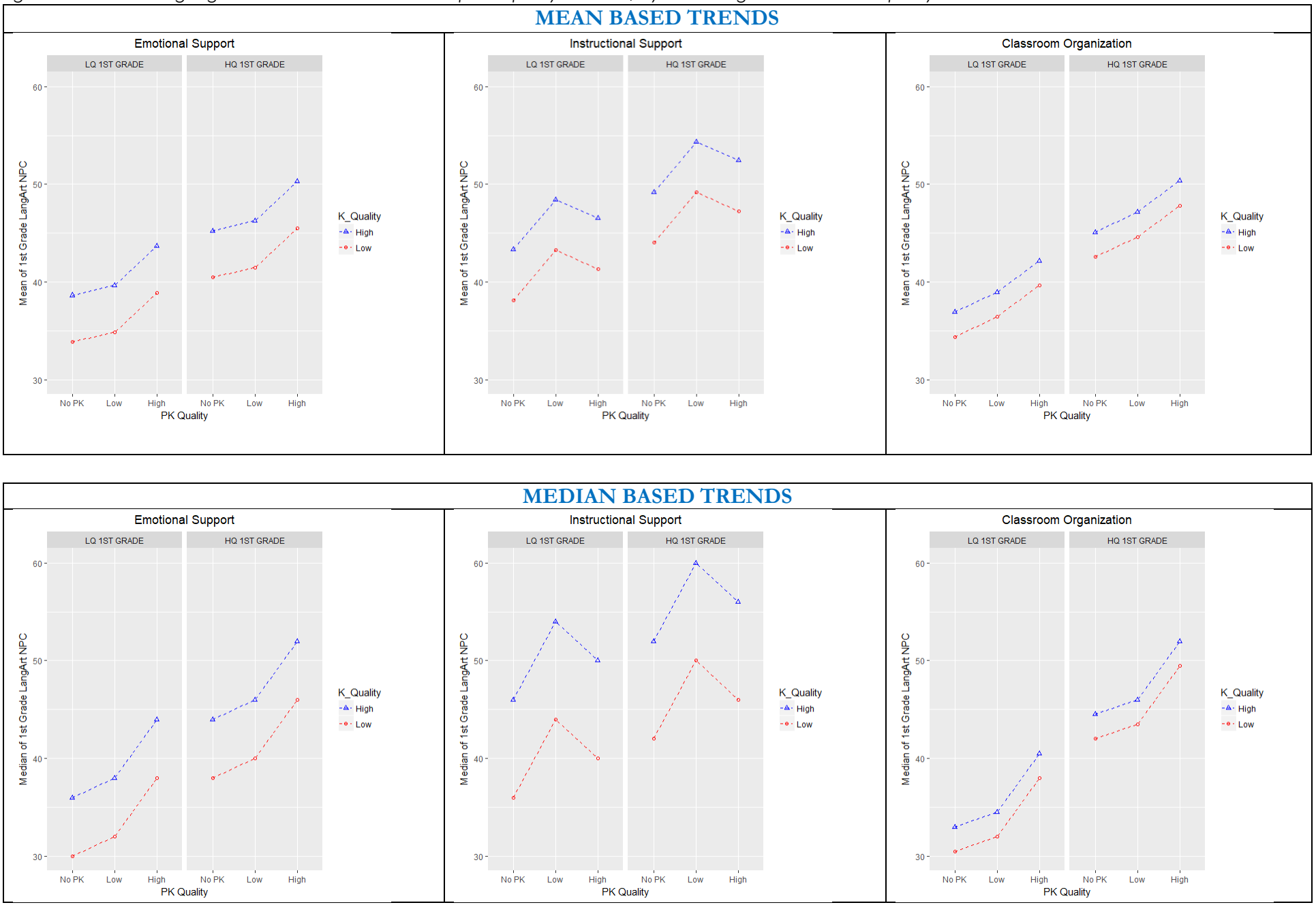
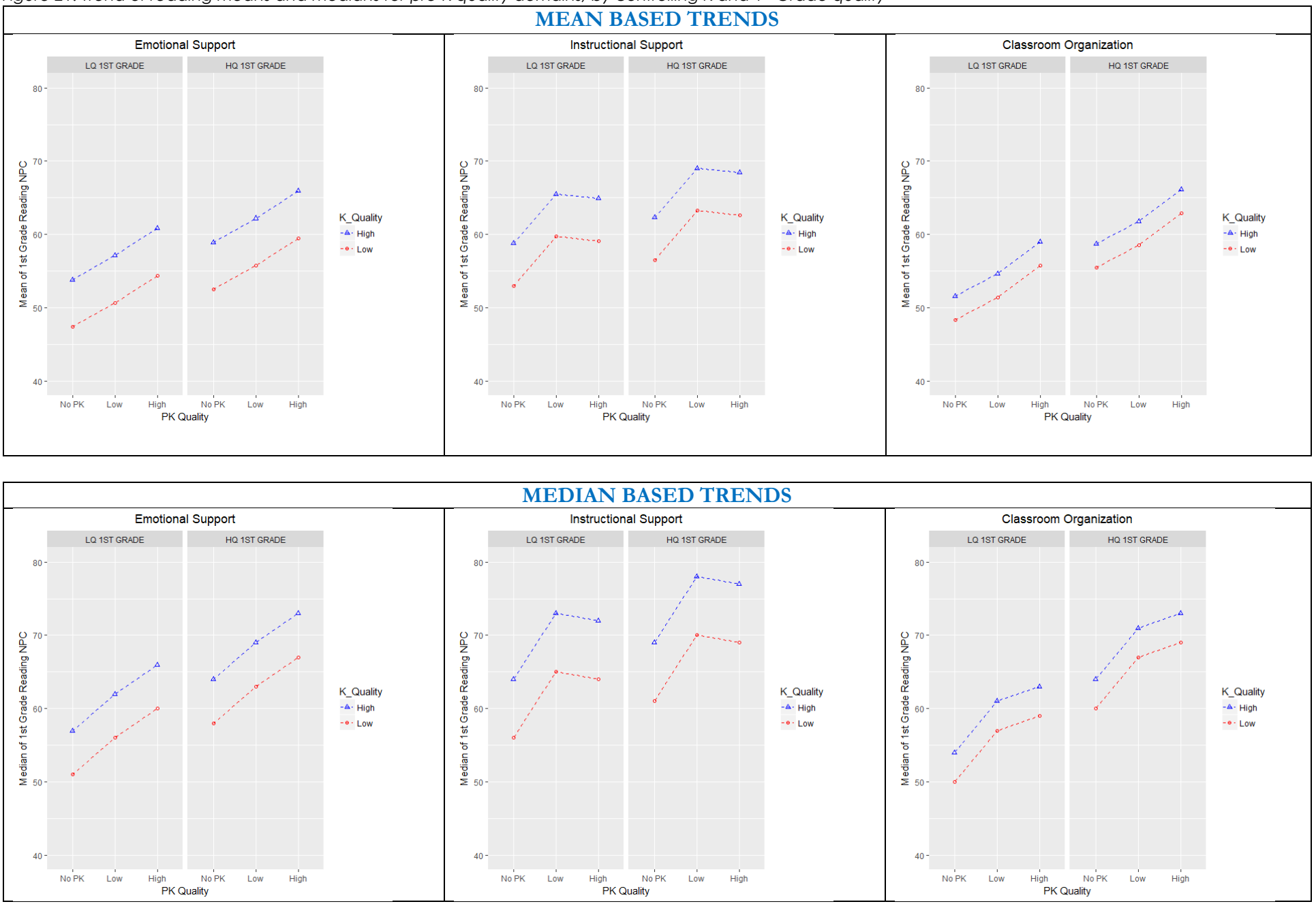


Figure 21. Trend of reading means and medians for pre-K quality domains, by controlling K and 1st Grade quality



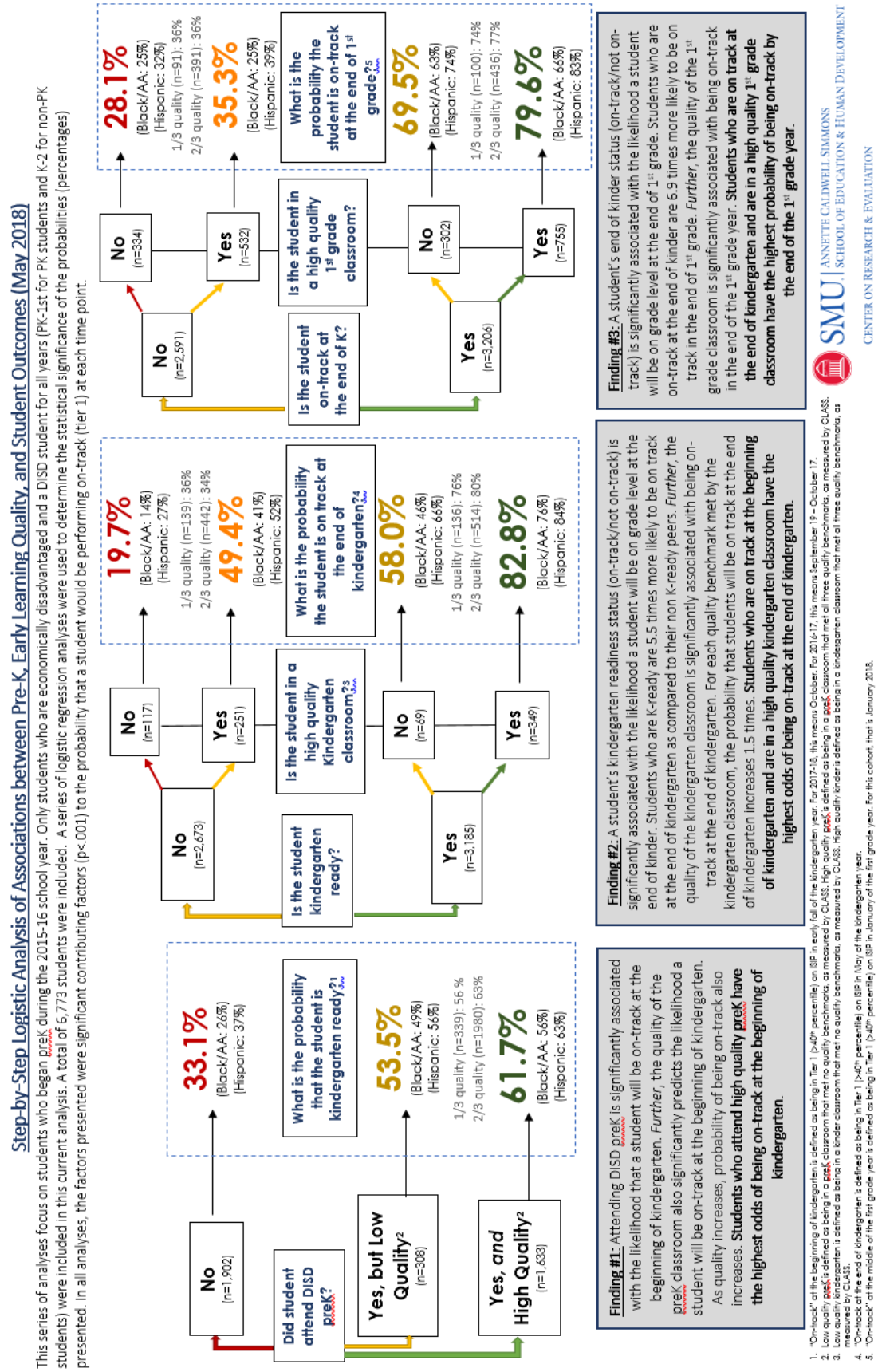
Limitations

- While the sample size is large compared to some available datasets, the cohort of students is relatively moderate in size and does not represent an entire grade-level within the district. For example, the students included in the "attended PreK in 2015-16" group totals 4,871 students. This is less than half of the 10,142 students who were enrolled in DallasISD PreK during that school year. Students who did not stay in the district through first grade, did not have a valid K readiness score at the beginning of the 2016-17 school year and who were not economically disadvantaged according to their demographic profile were omitted from the analyses. These cases are not missing at random from the dataset.
- CLASS® data is not available for all classrooms. It is available for most PreK classrooms and about 1/third of the district's K-2 classrooms. Therefore, the K and first grade analyses represents a sampling of district campuses and not the entire district.
- The ISIP assessment is not a comprehensive assessment of K readiness. While the assessment does measure early literacy skills that are predictive of later reading success or failure, a variety of critical cognitive, behavioral and emotional learning indicators that are also linked both to the CLASS® definition of early childhood quality and to outcomes throughout childhood and adulthood are not assessed.

Next Steps

- CORE plans to continue tracking this cohort of students (2015-16 PreK students) through third grade, which will be the 2019-20 school year. As additional outcome data becomes available, including end of second grade ISIP and TerraNova/Supera and, eventually, third grade STAAR, the long-term trajectories of these students and the effect of high quality early learning will be elaborated. Additionally, the same analyses will applied to other cohorts of DallasISD students.
- CORE and DallasISD seek to better understand other factors contributing to student outcomes and how CLASS® adds-to or interacts with these factors. An example might be exploring whether these pathways are more or less pronounced in specific feeder patterns or groups of campuses, whether they are influenced by student language, or specific dimension of the CLASS® framework. Other examples include exploring outcomes for different demographic subgroups of students.
- Future analyses will also consider the continuous nature of the CLASS® scores in addition to the current methods of considering the categorical nature of the scores. These findings would provide insight to the associations between gains in quality and gains in student outcomes.

Appendix A. Illustration of Analysis #1 Findings



[i] Phillips, D.A., Lipsey, M.A., Dodge, K.A., Haskins, R., Bassok, D., Burchinal, M.R., Duncan G.J., Dynarski, M., Magnuson, K.A., & Weiland, C. (2017). Puzzling it out; The current state of scientific knowledge on PreK effects. (https://childandfamilypolicy.duke.edu/wp-content/uploads/2017/04/PreKStudy_FINAL_ForWeb.pdf)
[ii] Ansari & Pianta (2018)
[iii] Duncan, G.J. & Magnuson, K. (2013). Investing in preschool programs. *The Journal of Economic Perspectives*, 27, 109-132.
[iv] Duncan & Magnuson 2013
[v] Lipsey, M.W., Farran, D.C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's academic achievement and behavior through third grade. *Early Childhood Research Quarterly.*
[vi] Bailey, D., Duncan, G.J., Odgers, C.L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness, 10, 7-39.*
[vii] Bassok, D., Gibbs, C.R., & Latham, S. (2018). Preschool and children's outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? *Child Development, 00,* 1-23.
[viii] Phillips, D., Gormley, W., & Anderson, S. (2016). The effects of Tulsa's CAP Head Start program on middle school academic outcomes and progress. *Developmental Psychology, 52,* 8, 1247-1261.
[ix] Pianta, R.C. & Hamre, B.K. (2009). Conceptualization, measurement and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38,* 109-119.
[x] Jones, S.M., Brown, J.L., & Aber, J.L. (2008). Classroom settings as targets of intervention and research. In M. Shinn, & H. Yoshikawa (Eds.), *Toward positive youth development:* Transforming schools and community programs (pp.58-77). New York: Oxford University Press.
[xi] Jones, S.M., Zaslow, M., Darling-Churchill, K.E., & Halle, T.G. (2016). Assessing early childhood social and emotional development: Key conceptual and measurement issues. *Journal of Applied Developmental Psychology, 45,* 42-48.
[xii] Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of the research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327-350.
[xiii] Farran, D.C. (2017). Characteristics of PreK programs that drive positive outcomes. (https://childandfamilypolicy.duke.edu/wp-content/uploads/2017/04/PreKStudy_FINAL_ForWeb.pdf)
[xiv] Phillips et al (2017).
[xv] Students who were *not* economically disadvantaged were small in number and filtered out of analyses.