Pre-Kindergarten Program Evaluation
Effectiveness of Pre-K and Kindergarten Readiness
Prepared by the Department of Research & Performance Management

August 2019

Author: Marie A. Sell, Ph.D.

Key Findings

Shelby County Schools (SCS) District staff requested that a program evaluation be conducted on the District's pre-Kindergarten (pre-K) program with the dual goals of: 1) providing information on the effectiveness of the pre-K program, and 2) identifying any bright spots, best practices, or other findings of note. SCS pre-K students and their teachers during the 2017–18 school year comprised the cohort for this evaluation. Data analyzed included student assessments during pre-K and Kindergarten and pre-K teacher observation scores.

The analyses of pre-K effectiveness found that:

- The percentage of pre-K students performing on grade level increased over 23 percentage points from December (54.7%) to May (78.0%) as measured by scoring in Tier 1 in *Istation Reading*.
- Students in Tier 2 or Tier 3 in December went from an average score of 11 points below the basal Tier 1 score to an average score of 3.3 points above the basal Tier 1 score in May.
- Of the 78% of pre-K students in Tier 1 in May, only 55% scored "Kindergarten ready" in reading on the NWEA MAP assessment at the beginning of Kindergarten. This discrepancy highlights the need for caution to not interpret formative assessment scores as if they are achievement assessment scores.

The analyses of bright spots and notable findings found that:

- There was little variability in the CLASS observation scores across the 176 pre-K teachers included in the analyses, especially in the Emotional Support domain and the Classroom Organization domain.
- The pre-K teachers' CLASS Observation scores had little impact on predicting students' Kindergarten readiness the following year.

This program evaluation was limited by difficulties in accessing pre-K data. Recommendations include focusing on pre-K data management before further pre-K program evaluations are conducted.

Shelby County Schools Pre-Kindergarten Program

In 2017–18, SCS supported 256 general education pre-K classrooms[1] at various locations throughout the District, making available 5,120 slots for general education pre-K students in a variety of settings. Over half (57%) the classrooms are located at school-based sites in which pre-K classrooms are housed in one of the District's elementary schools. (Nine of these classrooms are in the Ridgeway Early Learning Center, the District's only school designated exclusively for pre-K classrooms.) Thirty percent of the classrooms are located in area Head Start centers, and 13% are operated by the District's community partners and are located in child care centers or charter schools. This variety offers opportunities for families to select the setting that best matches their family's needs. For example, one family may prefer to drop off and pick up a pre-K child at the same elementary school in which they have older children enrolled. Another family may prefer a child care

[1] SCS Early Childhood also provides Head Start wrap-around services to 20 special needs classrooms operated through SCS Exceptional Children, bringing the total to 276 classrooms it supports. Data from students and teachers in the special needs classroom are not included in this program evaluation.

setting so they can couple the pre-K classroom experience with before- or after-school care in the same facility.

Despite the different settings for pre-K classrooms, the services provided to the students and families are relatively uniform. The length of the school day is the same for all classrooms. All classrooms follow the same pre-K curriculum: *Big Day for Pre-K.*[i] Each classroom has a teacher and a teacher's assistant. Teachers in school-based sites and at community partner sites must hold a Tennessee teacher license in early childhood education. Teachers at Head Start sites are not required to be licensed, but must have at least a Bachelor's degree in early childhood. Additionally, all sites use the same assessments to evaluate students and require teachers to be observed using the same observation tool. Given the consistency in student experiences across the various classroom settings, all analyses in this evaluation report examine the District's pre-K program overall.

The focus of this program evaluation is to examine the effectiveness of the District's pre-K program and to identify bright spots, best practices, or other notable findings. Of particular interest is to examine links between pre-K participation and students' academic readiness for Kindergarten.

Data Analyzed in Program Evaluation

To examine the first question about the effectiveness of the pre-K program, two sets of student data were analyzed: *Istation Reading* scores from pre-K and NWEA Measures of Academic Progress (NWEA MAP)[ii] scores from the beginning of Kindergarten. Istation is a computer-adaptive formative assessment that is administered to SCS pre-K students three times per year. Based on test scores, students are categorized into one of three instructional tiers. Tier 1 students are considered to be "performing on grade level." Tier 2 students are those who "are performing moderately below grade level and in need of intervention," while Tier 3 students "are performing seriously below grade level and in need of intensive intervention."[iii] The assessment considers the time of year when categorizing students' scores. That is, students must show more early literacy mastery as the year progresses to stay in the same tier as at the beginning of the year. Students who demonstrate more growth across the year could move up tier levels. Likewise, students whose scores stagnate or decrease over the year could move down tier levels. *Istation Reading* can help teachers identify students in need of reading interventions and track student progress across the school year. During the 2017–18 school year, pre-K students were assessed on Istation in December 2017, March 2018, and May 2018. Between assessment periods, all pre-K students worked on early literacy skills online using the *Istation Reading* program.

The second data set analyzed regarding the question of pre-K effectiveness is the NWEA MAP scores for the pre-K students taken at the beginning of their Kindergarten year. NWEA MAP is also a formative assessment that tracks student progress in reading and mathematics. Students receive RIT scores (scale scores) for each assessment along with a test percentile. In addition, data indicate whether students met their growth targets, and percentile rankings are provided for students' rates of growth. All SCS K-8 grade students take this assessment three times per year. Scores from the fall assessment for incoming Kindergarten students have been used by the District as an indicator of kindergarten readiness. During the 2018–19 school year, the fall NWEA MAP testing window for Kindergarten students was August 29 – September 19, 2018.

To identify program bright spots and notable findings, teacher observation data were used. In 2017–18, pre-K classroom teachers were observed and rated using the *Classroom Assessment Schools System (CLASS)*,[iv] which is an observation tool designed to focus on teacher interactions within the

classroom. Ten indicators combine to create three different dimensions of classroom interactions: Emotional Support – the use of teaching behaviors that help students develop positive social-emotional skills; Classroom Organization – the use of teaching behaviors that help students develop skills to gain the most from the learning environment; and Instructional Support – the use of teaching behaviors that help students develop language and cognitive skills. During the 2017–18 school year, the majority of the CLASS observations occurred during fall semester, and were conducted by Early Childhood instructional advisors.

Analyses and Results

Question 1: Effectiveness of the Pre-K Program

Analyses of Istation Data

Pre-K program effectiveness was first analyzed using a paired, pre-test/post-test score comparison of the December 2017 and May 2018 test scores. These two test times were selected (and the March 2018 assessment scores excluded) to be able to examine the greatest amount of program impact due to pre-K participation. To be included in the analyses, students had to have test scores for both the December 2017 and the May 2018 test administrations. Two thousand four hundred twenty-eight (2,428) students were included in the analyses.

The table below shows the percentage of students who scored in each tier level at each time period. In December, almost 55% of students were in Tier 1 or performing at grade level. Approximately 45% of students scored in Tier 2 or Tier 3, indicating some level of intervention was necessary. In May, the percentage of students in Tier 1 had increased to 78% and the percentages of students in both Tier 2 and Tier 3 had decreased. Overall, participating in the pre-K program boosted several students to perform on grade level in early literacy as measured by *Istation Reading*.

Percentage of Students in Each Istation Tier Level Students with Both Assessment Scores Only (N = 2,428)		
Tier Level	**December 2017**	**May 2018**
Tier 1	54.7%	78.0%
Tier 2	20.5%	12.5%
Tier 3	24.8%	9.5%

A second analysis of the effectiveness of the pre-K program examined student growth by analyzing scores of students who were in Tier 2 or Tier 3 in December (N = 1,101). Students' *Istation Reading* scores were subtracted from the lowest score required to be classified as scoring in Tier 1, thereby determining how many points each student was below Tier 1. In December, students were on average 11.0 points below the Tier 1 level. The same process was repeated for the same students in May, using May scores and tier cut points, at which time these students scored on average 3.3 points *above* the lowest Tier 1 score. Students' difference scores for December and May were analyzed using a paired t-test, which revealed a statistically significant difference ($t_{(1,100)} = 28.12$; $p < .01$). This increase showed that the Tier 2 and Tier 3 students, on average, gained enough early literacy skills to no longer be considered needing intervention, according to the Istation definitions.

Both analyses of Istation data from December to May revealed that the early literacy skills of students enrolled in the pre-K program increased over the year. Many students moved up tier levels, with approximately 23% more students scoring in Tier 1. Additionally, analyses of struggling students' scores revealed strong growth for this group over the year.

Analyses of Kindergarten Readiness

Another way to examine the effectiveness of the pre-K program is by examining how prepared pre-K students are when they enter Kindergarten the following school year. For these analyses, the same group of pre-K students' data were analyzed as were used in the previous analyses: students who have both December and May Istation data (N = 2,428). This restriction serves as a proxy to identify students who were enrolled in the program for at least six months (students who were substantially served). Since the pre-K program follows a continuous enrollment policy, students could be enrolled for only part of the year. Data for students enrolled fewer than six months may not reflect program impact due to a shortened participation time. Including these data would lead to inconclusive findings.

The District considers incoming Kindergarten students to be "Kindergarten ready" in reading and mathematics if they score at the 50th percentile or above on the fall NWEA MAP assessment. This equates to a MAP Reading RIT score (scale score) >= 141, and a MAP Mathematics RIT score >= 140. Approximately 75% of the students with both December and May Istation test scores also had NWEA MAP scores at the beginning of the Kindergarten year.

The first row of the table below shows that approximately 48% of the students who attended pre-K were Kindergarten ready in reading and approximately 42% in mathematics. However, it makes sense to limit these analyses to only pre-K students who were in Istation Tier 1 in May, or those who were "performing on grade level" at the end of pre-K. Students not in Tier 1, or those identified as needing some level of intervention, at the end of pre-K are unlikely to meet the District's criteria for Kindergarten Readiness. Again, approximately 75% of the May Tier 1 students had NWEA MAP scores. The second row in the table below shows that approximately 55% of students met the benchmark of being Kindergarten ready in reading and approximately 50% met the criterion in mathematics.

Percentage of Pre-K Students Kindergarten Ready on NWEA MAP				
	Reading		Mathematics	
	N	% KK Ready	N	% KK Ready
All pre-K students	1,852	48.3%	1,834	41.9%
May Tier 1 pre-K students	1,447	55.0%	1,433	49.3%

Discussion of Istation and NWEA MAP Results

Whether looking at all pre-K students or only those who were classified at Tier 1 students at the end of pre-K, roughly half the students were Kindergarten ready in reading at the beginning of the Kindergarten year. This is a notable drop from the pre-K year-end finding that 78% of students were performing on grade level (i.e., in Tier 1). This discrepancy between the end of pre-K and the beginning of Kindergarten for the same group of students could indicate summer learning loss. However, it also brings into question how closely aligned the pre-K and the early elementary school measures are that are being used by the District to assess students. Part of the difference has to do with the percentile cuts used as the benchmark criterion. Istation classifies scores in the 40th percentile or higher as Tier 1 scores.[vii] The District determined that NWEA MAP scores at the 50th percentile or higher in Reading met the benchmark of Kindergarten ready.

Another difference might be the specific content assessed by the two measures. Istation's subtests include Letter Knowledge, Listening Comprehension, and Vocabulary at the pre-K level. In

Kindergarten, NWEA MAP assesses four subtests: 1) Print Concepts, Phonological Awareness, Phonics; 2) Word and Sentence Composition, Writing; 3) Vocabulary Acquisition; and 4) Reading. While the category labels for NWEA MAP are broader than for *Istation Reading*, there is no way to know exactly what is being assessed. Individual assessment questions are not available for analysis for either of the measures. Therefore, it is not possible to know, for example, whether NWEA MAP captures more in the Print Concepts, Phonological Awareness, Phonics category in Kindergarten than Istation's Letter Knowledge category in pre-K. Even when the assessment category labels are relatively similar (e.g., Vocabulary and Vocabulary Acquisition) questions remain as to how similar the assessments are. For example, a vocabulary assessment asking students to correctly label pictures is quite different than an assessment that asks children for definitions of words. Again, access to the assessment content would be needed to answer these questions.

Finally, it is important to remember that both *Istation Reading* and NWEA MAP are formative assessments, not achievement tests. Both are designed to track student growth over time and to identify at-risk students for intervention. Tier 1 students in *Istation Reading* are defined as "performing on grade level." This does not mean that students in Tier 1 have mastered all the grade-level skills assessed. It simply means that based on the assessment, at that time, nothing flags the students for intervention. Therefore, students who are in Tier 1 at the end of pre-K are not guaranteed to score as "Kindergarten ready" at the beginning of Kindergarten.

Question 2: Bright Spots and Findings of Note

The second purpose of this program evaluation was to identify bright spots, best practices, or other findings of note. This question was examined by analyzing teacher observation data from the *Classroom Assessment Schools System (CLASS)* and examining whether those data are associated with Kindergarten readiness. The data file from the previous analyses on pre-K effectiveness was used for these analyses with an additional constraint. Any teachers whose rosters had only five (or fewer) students with complete data were also removed from the analyses. This was done in an attempt to limit the analyses to relatively stable classrooms. Any classrooms with more than five students with missing data were likely experiencing late enrollments and other disruptions for the majority of the students. As a result of these constraints, descriptive analyses of CLASS Observation data were conducted on 176 pre-K teachers.

CLASS rates teacher interactions within the classroom on three domains, all related to supporting positive student growth. The first domain, Emotional Support, is measured by indicators of Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. The second domain is Classroom Organization. It is measured by indicators for Behavior Management, Productivity, and Instructional Learning Formats. Finally, the Instructional Support domain is measured by indicators for Concept Development, Quality of Feedback, and Language Modeling.

Each indicator was rated from one to seven with higher scores indicating that more of the behaviors were observed. CLASS calculates domain scores by averaging the scores for the indicators within each domain.[2] The table below shows the mean rating score and the score range for each CLASS domain. The Emotional Support domain had a mean score of 6.2 out of a possible 7.0. The Classroom

[2] For the Negative Climate indicator in the Emotional Support domain, higher scores also indicate greater instances of negative climate interactions. The CLASS formula adjusts scores for this indicator when the Emotional Support domain score is calculated so that a higher score in Emotional Support indicates more instances of positive emotional support interactions.

Instruction domain had a mean score of 5.9, whereas the Instructional Support domain had a mean score of 4.1.

Teacher Scores on CLASS Observation Domains (N=176)		
CLASS Domain	**Mean Score**	**Score Range**
Emotional Support	6.2	4.0 - 7.0
Classroom Organization	5.9	3.5 - 7.0
Instructional Support	4.1	1.0 - 7.0

Predicting Kindergarten Readiness using CLASS Domain Scores

A set of multiple regression analyses was conducted to determine the relation between the CLASS observation scores the students' pre-K teachers received and students' NWEA MAP scores at the beginning of the kindergarten year. As discussed earlier, accessing data for the pre-K students was challenging. Therefore, much of the typical student-level data that are used as covariates or entered as predictors are not available for this data set. The data used as predictors in this analysis included the three CLASS observation domain scores, the student's *Istation Reading* May score and the student's age in months on the first day of kindergarten.

The first analysis was conducted using NWEA MAP reading scores as the outcome variable. Results indicated that students' *Istation Reading* May scores, students' age at the beginning of Kindergarten, the Emotional Support domain and the Classroom Organization domain significantly predicted NWEA MAP reading scores ($F_{(5,1836)} = 75.43$; $p < .001$), although combined they explained only 17% of the variance ($R^2 = .17$).

The Beta coefficients in the table below show the amount of impact any individual predictor has on the students' reading score. The students' *Istation Reading* score in May had the biggest impact. For every one standard deviation the *Istation Reading* score increased, the NWEA MAP reading score increased one-third of a standard deviation, or approximately 3.3 RIT scale score points. Students' age at Kindergarten, the Emotional Support domain observation score, and the Classroom Organization domain observation score all had lesser impact. Interestingly, Emotional Support had a negative impact on NWEA MAP reading scores (i.e., as Emotional Support scores increased, NWEA MAP reading scores decreased).

Beta Coefficients and Significance Levels for Multiple Regression Analysis of NWEA MAP Reading		
Predictor Variable	**Standardized Beta**	**Significance Level**
Istation Score in May	.33	<.001*
Age at Kindergarten	.18	<.001*
Emotional Support	-.12	.001*
Classroom Organization	.14	<.001*
Instructional Support	-.03	.31

*Denotes statistical significance at the .05 level.

The second analysis examined NWEA MAP mathematics scores as the outcome variable. The findings were quite similar to the findings for reading. Again, students' *Istation Reading* May scores, students' age at Kindergarten, the Emotional Support domain, and the Classroom Organization

domain all significantly predicted NWEA MAP mathematics scores ($F_{(5,1818)} = 86.79$; $p < .001$). Again, as with reading, combined these predictors explained very little (19%) of the variance ($R^2 = .19$).

The Beta coefficients presented below show that the students' *Istation Reading* score had the greatest impact on NWEA MAP mathematics scores. For every one standard deviation *Istation Reading* scores increased, NWEA MAP mathematics scores increased just over one-third of a standard deviation, or approximately 3.9 RIT scale score points. Age at Kindergarten had the next greatest impact, increasing NWEA MAP mathematics scores by one-fifth of a standard deviation for every one standard deviation students' age increased. This equated to approximately 2.2 RIT scale score points. The impact of the Emotional Support and Classroom Organization domains was minimal.

Beta Coefficients and Significance Levels for Multiple Regression Analysis of NWEA MAP Mathematics		
Predictor Variable	**Standardized Beta**	**Significance Level**
Istation Score in May	.36	<.001*
Age at Kindergarten	.20	<.001*
Emotional Support	-.07	.05*
Classroom Organization	.08	.04*
Instructional Support	-.04	.11

*Denotes statistical significance at the .05 level.

Predicting Kindergarten Readiness using CLASS Item Indicator Scores

There were two findings in the above analyses that led to another set of regression analyses. The first was that very little of the variance was explained using CLASS domain summary scores as predictors. The second was the finding that the Emotional Support domain was negatively associated with NWEA MAP scores at Kindergarten. The second set of regression analyses used the scores on the individual CLASS item indicators as predictors instead of using the three domain summary scores. Since the Instructional Support domain was not a significant predictor for either NWEA MAP reading scores or NWEA MAP mathematics scores, its underlying item indicators were not included in these regression analyses.

The Emotional Support Domain consists of:

- Positive Climate – such as social conversation, smiling, positive expectations, and respectful language
- Negative Climate – such as irritability, yelling, threats, and sarcasm
- Teacher Sensitivity – such as anticipating problems and planning accordingly, and providing individualized support
- Regard for Student Perspective – such as following students' leads, allowing choice, and encouraging student talk

The Classroom Organization Domain consists of:

- Behavior Management – such as clear expectations, efficient redirection, and consistency
- Productivity – such as few disruptions during learning time, clear instructions, and brief transitions

- Instructional Learning Formats – such as effective questioning, hands-on opportunities, and summaries of learning objectives

Again, two multiple regression analyses were performed. The first examined NWEA MAP reading scores as the outcome variable. The predictor variables included the seven item indicators delineated above along with *Istation Reading* May score and students' age at Kindergarten. The findings revealed that this model significantly predicted NWEA MAP reading scores ($F_{(9,1832)} = 43.37$; $p < .001$). The amount of variance explained by the equation (18%), however, remained quite low ($R^2 = .18$).

Students' *Istation Reading* scores and their age at Kindergarten, again, had a greater impact than the other predictors and the magnitude was the same as in the previous regression analysis. Only one item indicator from the Emotional Support domain and one from the Classroom Organization domain significantly impacted NWEA MAP reading score. Teacher Sensitivity from the Emotional Support domain negatively impacted NWEA MAP reading score. For every standard deviation the Teacher Sensitivity rating increased, NWEA MAP reading scores decreased by just over one-tenth of a standard deviation, which is approximately 1.2 RIT scale score points. Productivity, from the Classroom Organization domain, had approximately the same magnitude of impact, although it was in the positive direction.

Beta Coefficients and Significance Levels for NWEA MAP Reading Multiple Regression Analysis		
Predictor Variable	**Standardized Beta**	**Significance Level**
Istation Score in May	.33	<.001*
Age at Kindergarten	.18	<.001*
Positive Climate	.03	.51
Negative Climate	-.02	.37
Teacher Sensitivity	-.12	.001*
Regard for Student Perspective	-.03	.34
Behavior Management	.01	.87
Productivity	.11	.001*
Instructional Learning Formats	.01	.70

*Denotes statistical significance at the .05 level.

The multiple regression analysis for NWEA MAP mathematics scores was similar as well. The same set of predictor variables was entered into the equation which resulted in a statistically significant model predicting NWEA MAP mathematics scores ($F_{(9,1814)} = 49.80$; $p < .001$), which accounted for 20% of the variance ($R^2 = .20$).

The Beta coefficients presented in the table below indicate that *Istation Reading* scores and students' age at Kindergarten had the greatest impact on the model. The only item indicator to have a significant impact was Teacher Sensitivity, which again had a slight negative impact.

Beta Coefficients and Significance Levels for NWEA MAP Mathematics Multiple Regression Analysis		
Predictor Variable	**Standardized Beta**	**Significance Level**
Istation Score in May	.36	<.001*
Age at Kindergarten	.20	<.001*

Positive Climate	.05	.17
Negative Climate	-.04	.10
Teacher Sensitivity	-.09	.02*
Regard for Student Perspective	-.04	.18
Behavior Management	-.02	.68
Productivity	.06	.09
Instructional Learning Formats	.003	.91

*Denotes statistical significance at the .05 level.

Discussion of CLASS Observation Analyses

Several interesting findings emerged with the second set of multiple regression analysis. First is that only one item indicator from each CLASS domain significantly impacted the outcome variables. For the Emotional Support domain, it was Teacher Sensitivity. For Classroom Organization, it was Productivity. Although other item indicators make up the overall domains, when entered as separate predictors they did not impact the model. This is not to say that the other item indicators are insignificant in the context of a pre-K classroom, but rather they did not have a statistically significant impact on the outcome variables examined in these analyses (NWEA MAP scores at the beginning of Kindergarten).

A second interesting finding is that, although it had a very minor impact, Teacher Sensitivity consistently entered in the negative direction. As teacher sensitivity ratings increased, NWEA MAP scores decreased. On the surface this seems puzzling. However, a close look at the kinds of teacher actions rated by this indicator shows it partially measures teachers' ability to anticipate and resolve potential problems. Too much of this on the teacher's part might inadvertently be negatively impacting student cognitive growth. Minor problems that occur in classroom settings often allow students to develop problem solving skills. For example, there may not be enough art supplies on the table for each student to do the art project. Children without supplies have to figure out how to resolve the problem (e.g., by asking the teacher for more supplies, or by negotiating how to share the supplies among the students). A vigilant teacher who always has enough art supplies on the table to prevent potential struggles among students, may also be removing some naturally-occurring problem-solving opportunities which help develop cognitive skills in children.

Finally, it is important to note that the total amount of variance explained by any of the regression models ranged from 17-20%, with the majority of this coming from students' Istation May scores and their age at Kindergarten. Although the teacher interactions rated in the CLASS Observation rubric do not appear to align with NWEA MAP scores in Kindergarten, increasing and strengthening these kinds of teacher interactions do help with young students' social-emotional learning and likely create a more positive classroom climate for students.[v] Ultimately, though, factors other than those used in the regression analyses accounted for 80% or more of students' NWEA MAP scores. While identifying what these factors are is an interesting question to pursue, it is beyond the scope of this report given the limited data available for analysis.

Discussion and Recommendations

Lack of Accessible Data

The purpose of this program evaluation of the District's pre-K program was two-fold. First, it sought to evaluate the program's effectiveness; and second to identify bright spots or best practices. Unfortunately, both these efforts were hampered by lack of accessible data for pre-K students.

Students in the pre-K program are assessed with multiple instruments to track their progress. At pre-K screening, which determines whether a child is eligible to enroll in the program, children are assessed with the *Brigance Early Childhood Screen III (3-5 years)*[vi], which measures their developmental level in the areas of Physical Development, Language Development, Academic Skills/Cognitive Development, and Self-Help and Social-Emotional Skills. (In addition to the developmental screening, family income information and data on other academic risk factors are collected at this time.) Once enrolled in pre-K, students are assessed with the *Brigance Inventory of Early Development*[vii] three times per year to track skill mastery in the domains listed above plus Adaptive Behavior (daily life skills). Additionally, their receptive vocabulary skills (words they understand) are measured using the *Peabody Picture Vocabulary Test – 4 (PPVT-4)*[viii] at the beginning and the end of the school year. Finally, to track progress in early literacy skills, students are assessed three times per year using *Istation Reading,*[ix] which is a formative assessment.

Scores on all these measures would provide a rich data set to examine developmental and academic gains over time. However, due to technical challenges with extracting student assessment data from the pre-K student data platform that was used in 2017–18 (Child Plus), very little of this information is available for analysis. The only pre-K data available to be analyzed in the evaluation were student scores in *Istation Reading*.

The analyses of *Istation Reading* scores did reveal that students' early literacy skills generally improved over the year, however it would be expected that students' skills improve as they work through the Istation curriculum online. It would be helpful to be able to corroborate this finding by analyzing additional pre-K data related to academic progress. For example, pre-K students are assessed twice a year with the *Peabody Picture Vocabulary Test – 4 (PPVT-4)*, which is a standardized receptive vocabulary assessment appropriate for ages 2-1/2 years through adulthood. Standard scores from this assessment can be compared to determine vocabulary growth that is independent from a child's age. Therefore, increases in vocabulary identified by this assessment could speak more directly to the impact of pre-K participation on students' academic progress. Additionally, since students' academic progress is influenced by outside factors, it would be helpful to be able to access data on risk factors that may affect learning. In light of the data challenges involved in this program evaluation, the following recommendation is made:

Recommendation 1 – Before undergoing another program evaluation of the District's pre-K program, a system needs to be in place to assure that the academic and non-academic data of the pre-K students can be accessed in a format that can be analyzed. This recommendation is two-fold. First, the District and vendor must establish that student-level data files are easily accessible and available to the District from any data platform being considered for assessments or other student tracking.

Second, measures should be in place to ensure that the appropriate data formats are pulled for analyses. Again, using the *PPVT-4* as an example, students' scores are reported in a number of different ways, including a raw score, a standard score, and an age equivalent. The different score

formats provide different kinds of information about the students' receptive vocabulary levels and choosing the correct score to use would depend on the analysis to be conducted. It is strongly recommended that Early Childhood hire a data analyst who is familiar with both using student data in the classroom *and* analyzing pre-K assessments to help manage the data collection process of pre-K student data. This data analyst could also help coordinate data access for future program evaluations.

Update from Early Childhood Program: Use of Assessment Data –As of the 2019-20 school year, the pre-K program will no longer be administering the PPVT-4. The Brigance Screening and Brigance Inventory assessments will continue to be used. The screening process will begin during pre-K enrollment and continue throughout as students acclimate to the learning environment. Parent input for the behavioral screening will be gathered during home visits and parent-teacher conferences. Data from these research-based measures will be used to develop individual cognitive and behavioral growth goals for each student enrolled in pre-K.

The pre-K program will continue to use Istation three times per year. The data from these assessments will be used to inform program plans and decisions, including decisions on individualization, ongoing monitoring, and resource allocation, and to inform and influence teacher practice through training, technical assistance/coaching, and professional development.

Revisit CLASS Observations

A second factor that emerged during the data analyses was how little variation there was in the CLASS observation scores across the 176 teachers whose observation data were included in the analyses. Possible scores for each domain ranged from 1.0 to 7.0, with a score of 1.0 indicating that few of the interactions in the domain were observed and 7.0 indicating that several were observed. The Emotional Support domain had an average score of 6.2, with scores ranging from 4.0 – 7.0. The average score for the Classroom Organization domain was 5.9 with scores ranging from 3.5 – 7.0. Both these domains had a standard deviation of less than 1.0. The Instructional Support domain had a bit more variability in scores, with an average of 4.1 and a range from 1.0 – 7.0. The standard deviation for this domain was just above 1.1.

The tightness of the scores for 176 observations could signal that the observers are experiencing observation drift. Observation drift is a phenomenon that occurs when trained observers conduct multiple observations. Over time, their judgments about whether observed interactions meet the criteria to be counted begin to drift from the observation scale standards. Many entities that use classroom observations as part of their ratings (such as Tennessee's Star Quality Program[x] for rating child care centers in the state), require trained observers to undergo reliability checks periodically. Observers whose scores are no longer reliable are re-trained so they can provide accurate observation scores.

Recommendation 2 – Personnel who conduct the CLASS Observations in the pre-K classrooms should be re-trained periodically to ensure they continue to be reliable observers of the classroom interactions. Further, it is recommended that the District not rely on train-the-trainer models to train new observers or re-train experienced observers. Instead staff members should attend sessions conducted by CLASS Observation trainers.

Update from Early Childhood Program: CLASS Observation Training – Beginning in the 2019-20 school year, CLASS Observation trainings will be aligned with SCS-specific classroom performance

expectations. Professional development for classroom teachers will be provided by trained SCS pre-K staff focusing on high-quality student-teacher interactions that are expected to lead to improved academic and social-emotional outcomes for students. Professional development support will provide teachers with engagement and instructional strategies that will impact learning, development, and lifelong achievement.

Realistic Expectations of Assessments

A third issue related to a pre-K program evaluation would be to ensure that there are realistic expectations about the information that can be gained given the data and assessment scores being analyzed. For example, individual assessment questions in *Istation Reading* and NWEA MAP are not available so District staff cannot know the extent to which these assessments align either in the content of what is assessed (outside of broad domains) or in how to interpret test scores as they relate to each other. This will also be a consideration when the District switches from NWEA to Illuminate Education[xi] as the provider of its universal screener. At the very least, District staff should remember that formative assessments and universal screeners are not achievement tests, and to consider them as such could provide a false reading of program effectiveness.

Recommendation 3 – Before engaging in another full-fledged evaluation of the pre-K program, preliminary and exploratory data analyses should be conducted to determine how pre-K *Istation Reading* scores align with Illuminate or scores from other new Kindergarten measures. SCS will have the opportunity to determine relationships between these two assessments during the 2019-20 school year, since pre-K students will complete the Illuminate assessment for the first time in spring 2020.

Update from Early Childhood Program: Use of Illuminate Screenings – The spring Illuminate/ Fastbridge administered by the pre-K teachers will identify each child's instructional and foundational levels of readiness for Kindergarten. In fall 2020, Kindergarten teachers will compare students Kindergarten Illuminate assessments to their pre-K assessment from the previous spring to quantify each student's performance level. In addition to providing Kindergarten teachers with instructional benchmark data, this comparison can also speak to summer learning loss from pre-K to Kindergarten entry.

Examine Non-academic Impact of Pre-K Participation

Finally, there are many anecdotal accounts from Kindergarten teachers that students who have attended pre-K are much better prepared for Kindergarten in non-academic areas that are important to the day-to-day activities within a Kindergarten classroom. For example, pre-K participants are already familiar with school culture such as having a daily structure and routines, and with classroom expectations such as requesting permission and transitioning from one activity to another. According to teachers, students who already know "how school works" when they come to Kindergarten are able to settle in more quickly and begin focusing on academics compared to peers without this prior knowledge.

Recommendation 4 – A pre-K evaluation should be conducted to analyze the non-academic benefits of attending pre-K, including how familiarity with school routines is associated with being ready to focus on academic instruction. A non-academic analysis could potentially also examine Kindergarten students' family behaviors as well to determine whether children's participation in pre-K is associated with factors in the family's control, such as attendance or tardy rates in Kindergarten, or establishing routines for before-care and aftercare.

Endnotes

[i] *Big Day for Pre-K* can be accessed at https://www.hmhco.com/products/big-day-pre-k/
[ii] *NWEA Measures of Academic Progress* can be accessed at https://www.nwea.org/map-growth/
[iii] Mathes, P., Torgesen, J., & Herron, J. (2016). Istation's Indicators of Progress (ISIP) Early Reading Technical Report, pg 5-4. Accessed on May 14, 2019 at https://www.istation.com/Content/downloads/studies/er_technical_report.pdf
[iv] The *Classroom Assessment Scoring System (CLASS)* can be accessed at https://teachstone.com/class/
[v] Otrosky, M. M., and Yung, E. Y. (no date). *Building Positive Teacher-Child Relationships*. A *What Works Brief* developed by the Center on the Social and Emotional Foundations for Early Learning accessed at http://csefel.vanderbilt.edu/briefs/wwb12.pdf
[vi] The *Brigance Early Childhood Screen III (3-5 years)* can be accessed at https://www.curriculumassociates.com/products/brigance/early-childhood
[vii] The *Brigance inventory of Early Development* can be accessed at https://www.curriculumassociates.com/products/brigance/early-childhood?GTM_ProductCard
[viii] The *Peabody Picture Vocabulary Test – 4* can be accessed at https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Brief/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html
[ix] *Istation Reading* can be accessed at https://www.istation.com/Reading
[x] Notes on assessor training for the Tennessee Star Quality Program can be accessed at https://starquality.sworpswebapp.sworps.utk.edu/program-assessments/
[xi] Information in Illuminate Education assessments can be accessed at https://www.illuminateed.com/