2015 Educational Return on Investment Report

2013-2014 Program Evaluations

March 2015

Office of Accountability
Department of Research, Evaluation, and Assessment

This page left intentionally blank.

Primary Author

Reem Abdelrazek
Data Analyst

Contributing Authors

John Beckett
Director

Clint Sattler
Supervisor

Steve Rudder
Senior Data Analyst

Marie Lunde
Research Specialist

About the Department

The Department of Research, Evaluation, and Assessment (REA) is a multi-faceted team that serves the district within the Office of Accountability. The REA department is comprised of the Supervisor of Research and Evaluation, the Supervisor of Assessment, a senior data analyst, a data analyst, and two specialists. The department is responsible for state accountability measures, administration of all district-wide assessments, program evaluation, researching curricular data, communicating data to appropriate stakeholders across the district, and providing its analytical expertise to assist school leaders in making student-centered, data-driven decisions. In addition to these responsibilities, the REA team also serves as the gateway for external organizations requesting access to data from the Knox County Schools to include in third-party research.

About the Office of Accountability

The Office of Accountability operates under the leadership of the Chief Accountability Officer. The office is responsible for district accountability and organizational performance, with the ultimate goal of increasing student academic achievement. Staff members lead efforts to interpret data, identify root causes, and provide actionable feedback to inform strategic planning and resource allocation. The Office of Accountability directs and coordinates the following areas: Elementary and Secondary Education Act compliance; assessment; research; program evaluation; performance evaluation data collection and support; performance-based compensation data collection and support; federal programs; strategic planning and improvement; and competitive grant funding and management.

Contributors

Knox County Schools

Dr. Jim McIntyre
Superintendent

Curriculum & Instruction

Dr. Elizabeth Alves
Assistant Superintendent, Chief Academic Officer

Dr. Clifford Davis
Executive Director, Secondary Education

Susan Turner
Executive Director, Elementary Education

Melissa Massie
Executive Director, Student Support Services

Millicent Smith
Executive Director, Curriculum, Instruction, & Professional Development

Theresa Nixon
Director, Instructional Technology

Human Resources

Dr. Kathy Sims
Chief Human Resources Officer

Dr. Rodney Russell
Director, Director of Human Capital Strategy

Public Affairs Office

Melissa Ogden
Director, Public Affairs

Office of Accountability

John Beckett
Director, Research and Evaluation

Laurie Driver
Supervisor, Assessment

Clint Sattler
Supervisor, Research and Evaluation

Steve Rudder
Senior Data Analyst

Reem Abdelrazek
Data Analyst

Beth Boston
Assessment Specialist

Marie Lunde
Research Specialist

Ginnae Harley
Director, Federal Programs

Keith Wilson
Director, TAP Program

Christy Hendler
Director, Planning and Improvement

Jamie Wilson
Grant Development Manager

Finance Office

Ron McPherson
Executive Director, Finance

Lizabeth McLeod
Director, Budget

External Organizations

Great Schools Partnership

Contents

Glossary of Frequently Used Terms

APEX Advance, Perform, Excel. Knox County Schools' strategic compensation system introduced in school year 2011-2012.

CBM Curriculum-Based Measurement. KCS uses AIMSweb as its universal screener to monitor student progress in literacy and numeracy based upon CBM.

EOC End-of-Course exam. EOC exams are state-mandated assessments for English I, II, and III; Algebra I and II; Biology I; Chemistry I; and U.S. History.

EROI Educational Return on Investment.

FY Fiscal Year. The investment analyses in this report refer to fiscal year 2014 (FY14) expenditures.

KCS Knox County Schools. The KCS is the third largest school district in Tennessee. KCS serves 58,000 students.

ILC Individual Learning Cycle. ILCs are personalized professional development and support for teachers in collaboration with instructional coaches.

NCE Normal Curve Equivalent. NCEs are the unit of measurement used to refer to student comparative performance on state assessments in grades 4 – 8. While percentiles are bunched at the mean under a normal curve, NCEs maintain equal length intervals.

PLC Professional Learning Communities. PLCs are collaborative planning sessions based on the model created by Richard and Rebecca DuFour.

REA Department of Research, Evaluation, and Assessment (Knox County Schools).

RLA Reading and Language Arts. RLA is a specific subject assessed by the Tennessee Department of Education.

RTI^2 Response to Instruction and Intervention. A statewide initiative led by the Tennessee Department of Education that is based on a three-tier framework. RTI^2 promotes recommended practices for integrated general and special education for students.

SAT-10 Stanford Achievement Test Series 10 (also known as K – 2 Assessment). The SAT 10 is a norm-referenced assessment utilized in KCS for students in Kindergarten through grade 2.

STC School Technology Challenge. This refers to a pilot program implemented in the KCS in SY1314 in which 11 schools received technology grants to provide personalized learning via one device per one student.

SY School Year. This EROI Report evaluates the 2013-2014 school year, SY1314.

TAP The System for Teacher and Student Advancement. A school reform model developed by the National Institute for Excellence in Teaching (NIET), TAP provides teachers with career advancement opportunities, job-embedded professional development, and performance-based compensation.

TCAP Tennessee Comprehensive Assessment Program. The TCAP exams are those administered by the Tennessee Department of Education in grades 3 – 12 to assess student mastery of the state standards.

TEAM Tennessee Educator Acceleration Model. TEAM is the annual evaluation process for all school-based certified staff as required by Tennessee state statute.

TVAAS Tennessee Value-Added Assessment System. TVAAS is a statistical model that seeks to measure the impact of teachers, schools, and districts on student academic growth. The Tennessee Department of Education contracts with the SAS Institute to complete the TVAAS calculations.

Executive Summary

Continuing an important practice the Office of Accountability began in 2012, we present the third annual Educational Return on Investment (EROI) Report, which provides an in-depth analysis of various programs of the 2013-2014 school year (SY1314) in the Knox County Schools (KCS). The Office of Accountability is committed to providing accurate, objective, relevant, and timely evaluations of district initiatives to truly measure the return on our educational investments and to advance a student-centered, data-driven culture.

It is important to assess the success of district programs for several reasons. Resources continue to be spread thin across the district. As part of the district's strategic vision, the Office of Accountability provides KCS stakeholders with information as to what is working, what needs adjusting, and what needs ceasing. More important, we must know which programs are helping us to achieve excellence for all children. Our strategic initiatives should ultimately result in successes for our students and our larger community.

The following report is based on program evaluation and analysis conducted by the Department of Research, Evaluation, and Assessment (REA), a department within the Office of Accountability. The qualitative and quantitative data components were captured in several ways by the REA staff with the assistance of the Curriculum and Instruction department. Recommendations were also made in concert with program leaders and stakeholders.

What's Inside?

The 2015 EROI Report includes three sections that present an increasing depth of analysis and detail.

1) The **executive summary** outlines the programs evaluated and the most compelling themes and considerations that emerged from the evaluations.
2) The **management reports** provide summary information about the programs, investment analyses, major findings, and recommendations.
3) The **technical reports** contain a detailed evaluation process for each program in terms of methodologies, statistical analyses, results, and considerations.

Every program evaluated in the EROI Report stems from the goals of the Knox County Schools Strategic Plan. The district adopted a five-year strategic plan, *Excellence for All Children*, in 2009. Although the district has now adopted a new strategic plan that spans 2015 – 2019, the SY1314 programs we evaluated fell under *Excellence for All Children*. The initiatives included in 2015 EROI report include the following:

Initiative	Description
Family & Community Engagement	This initiative is comprised of the Parents as Teachers and Community Schools program, under goal three of the 2009 strategic plan, Engaged Parents and Community. Parents as Teachers is a community project aimed at empowering parents to be their child's first teacher. We evaluated one aspect of Kindergarten readiness of the students whose parents participated in the Parents as Teachers program. Community Schools is a program that offers expanded after-school services in partnership with public agencies and non-profit providers. Our review of Community Schools analyzed the impact of participation in the Community Schools program on student attendance, behavior, and academic growth.
Personalization & Differentiation	This initiative involves literacy intervention and one-to-one personal technology devices. These programs fall under both goals one and four of the 2009 strategic plan, Focus on the Student and Infrastructure: Enabling Student Learning, respectively. The literacy intervention program is a district-wide initiative to increase comprehension and fluency in early grades, and is comprised of the materials, support, and personnel. We evaluated the impact of the early literacy intervention program. The additional elementary reading support review centered on instructional assistants hired specifically to provide intervention services. The literacy analyses concentrated on how these initiatives impacted student growth on SAT-10 and TCAP assessments. The School Technology Challenge is a program meant to fully integrate technology with the main objective of increasing teacher effectiveness to drive increases in student outcome data. The evaluation focused on student perceptions, outcomes, and teacher performance metrics.
Teacher Support	This initiative encompasses the work of instructional coaches and lead teachers as part of goal two from the strategic plan, Effective Educators. Instructional coaches support teachers through individual learning cycles and professional learning communities. The evaluation focused on observation and TVAAS results for teachers receiving coaching support in ILCs. We also reported on several points of qualitative data regarding PLCs. Lead teachers support instruction through the TEAM observation process. The lead teacher review highlighted the amount of observations conducted by lead teachers in TEAM schools.
Strategic Compensation	This initiative relates to incentivizing instructional excellence via the TAP evaluation model and the APEX reward system as part of goal two from the strategic plan, Effective Educators. TAP is a comprehensive school reform model with the goal of increasing teacher recruitment, retention, motivation, practices, and performance. The data reviewed as part of the TAP evaluation included teacher observation scores, perceptions, and retention rates, and student outcome data. APEX is the district strategic compensation system for TEAM schools. The evaluation of APEX contained an analysis of teacher job satisfaction, leadership opportunities, observation and TVAAS scores, and retention.
Staffing Models	This initiative is focused on the logistical nature of schedules and staffing in secondary schools as part of goal four from the strategic plan, Infrastructure: Enabling Student Learning. As part of the staffing ratio analysis, the relationship between class size and student performance was evaluated. The block scheduling evaluation compared the impact on high school staffing allocations using block versus six-period scheduling.

Bright Spots

Each program evaluation uncovered positive and negative results. We want to call attention to the successes so that district practitioners can learn from these "bright spots." District leaders and stakeholders are encouraged to read the management reports for each program evaluated for the full scope of the findings and recommendations.

- **Parents as Teachers**: This is a community program that reached over 240 young students and their families. Students for whom we had testing data showed directionally higher results on the early literacy assessment when compared to demographically similar students whose parents did not participate in the program.
- **Community Schools**: Overall, Community School student absences decreased compared to their peers. There was also growth in academic achievement for Community School students overall.
    - Community School students at Christenberry and Pond Gap Elementary schools showed growth in Reading, Language Arts, and Math proficiency.
- **Additional Elementary Reading Support**: Our evaluation showed that there were several pockets of success on a grade-by-grade level.
    - AERS students at Green and Gibbs Elementary schools in particular showed growth in three out of five grades when compared to the rest of the school.

- **School Technology Challenge**: In its first year of implementation, the STC pilot demonstrated positive gains.
    - There has been an increase in student engagement as captured by student perception surveys as well as teacher focus groups.
    - There was directional evidence of increases in school-level TVAAS.
    - STC teachers generally felt they were better prepared to deepen personalization this school year (SY1415).
- **Instructional Coaches**: Instructional coaches generally have the desired effect on struggling teachers and promote data-driven practices in our schools.
    - KCS teachers who earned the lowest scores prior to participating in learning cycles with coaches showed significant growth on both their observation scores and TVAAS gains after participation in ILCs.
    - Survey data suggested that positive perception around the intended purpose of PLCs increased among respondents since the revamp of the coaching model in 2012. These results suggest that educators perceive that PLCs are used, and used properly.
- **TAP**: Our evaluation of the TAP system showed some positive growth in high schools.
    - TAP teachers exhibited higher rates of change in observation scores in their first four years of teaching. This is evidence that the TAP system provides novice teachers with a better understanding of the classroom observation rubric, or that TAP is recruiting better prepared teachers.
    - TAP high schools exhibited increased mean TVAAS gains when compared to their TEAM counterparts.
- **APEX**: There were few but tangible increases in outcome data.
    - TEAM staff assumed more leadership responsibilities as measured on the APEX leadership rubric, evidence that APEX is positively affecting the number of teachers that are assuming leadership responsibilities in the district.

By highlighting where things are working, the district can continue to build best practices around program implementation and evaluation. Developing standards of practice based on these successes can accelerate the district's ability to expand upon those successes.

Operational Themes

In addition to discovering how these district initiatives are faring, several operational themes emerged from our return on investment analyses.

- **Our district must continue to be strategic about who it places in intervention and support programs.** The intended participants of an intervention, whether teachers or students, matter. We have repeatedly found negative effects on those who participated in an intervention for which they did not qualify. The converse has also been true, those who qualified for and participated in an intervention showed growth.
    - Our evaluation showed that students who participated in literacy intervention—but who did not qualify for intervention based on two metrics—often scored below their predicted performance.
    - Teachers who were below expectations prior to participating in individual learning cycles (ILCs) showed significant growth on their observation scores and TVAAS gains after ILC participation. Teachers who were at or above expectations tended to perform worse on their observation scores and TVAAS gains after ILC participation.
- **Collaboration is an important component of the district's work.** As we have reported before, and as evident in the district's continued efforts, collaboration among stakeholders benefits our students.
    - The ongoing collaboration with the Great Schools Partnership, the University of Tennessee, and other local organizations helped the district deliver quality services to our students and families.
    - Collaboration of teachers and instructional coaches resulted in growth for those teachers who most needed instructional support.
- **Although our data quality is improving overall, data collection and related processes continue to be a challenge in our district.** While KCS strives to fully embrace a data-driven culture—and our central data systems are aiding in that pursuit—data collection practices differ by building. These differing practices create varying ranks of data quality, which can ultimately compromise program evaluations. The district should provide best practices and training around appropriate data entry into central data warehouses.
    - In the 2014 EROI, we reported that tracking Voyager students did not capture all the necessary data. The data collection form was updated to address those concerns, and as such, tracking students who participate in Voyager-based intervention programs improved. Incidentally, the data from VPORT, the vendor tracking system, was less reliable and difficult to use to corroborate school participation rolls.

  - Scheduling data used for both our staffing ratios and block scheduling analyses required manual screening. While Aspen is a good central resource, the manner in which schools enter certain scheduling data can be inconsistent and ambiguous.
- **Continuous improvement and fidelity of implementation continue to be opportunities for improvement.** With a district the size of Knox County Schools, it is understandably a massive undertaking to monitor all of the district's initiatives. Nonetheless, since the release of our first return on investment report in 2012 that brought light to this issue, the district has made great strides toward purposeful implementation with fidelity.
  - When comparing schools that implement the same program, the question of fidelity arises when there are gains in select schools while other participating schools showed negative growth.
  - Evaluation efforts should be considered when designing district programs, which will allow us to link program activities to measureable outcomes.
    - Assuring that program evaluations are based on accurate metrics requires that program leaders develop appropriate metrics that take into consideration data sources, data quality, and timeliness of implementation and any associated assessments.

Management Reports

The following section contains management reports of the programs the REA evaluated. These reports offer information about the programs, a brief investment analysis, and the findings and recommendations related to each program evaluation. These management reports are not technical and do not provide the details of any statistical analysis. Additional data about methodology or specific results can be found in the next section, Technical Reports.

Family and Community Engagement

KCS understands and appreciates that teachers and schools alone cannot do the work of supporting and educating children. Families, neighborhoods, and communities are integral to a student's success. District leadership specifically set out a goal to engage parents and community stakeholders to help achieve excellence for every child.

Engaged Parents & Community

- This goal concerns the role of parents and the community in achieving our vision.
- KCS acknowledges that these stakeholders are important players in supporting the education of all students and that we must build family educational efficacy and coordinate focused district partnerships in order to make that happen.
- Collaborative partnerships with the community will help us achieve the strategic initiatives set out in this plan.

Source: Knox County Schools Strategic Plan (2009-2014)

Within that framework, KCS works with community organizations and the University of Tennessee on several fronts, but most notably, the Parents as Teachers and Community Schools programs. These are two school-based programs that engage pre-K to grade 5 students to promote active parenting, positive family and school relationships, and holistic support of students at these schools.

Investment Analysis

Funding for the Family and Community Engagement initiatives came from community organizations. No general purpose funds from the Knox County Schools were spent toward either program in fiscal year 2014 (FY14). As such, we do not have a KCS-based cost per student to report.

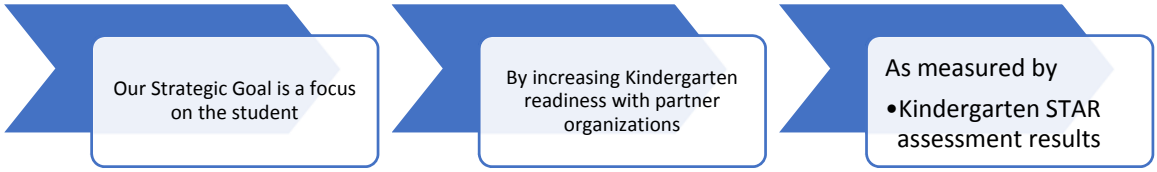
PARENTS AS TEACHERS

The Parents as Teachers program is a joint effort between Knox County Schools (KCS) and the Great Schools Partnership of Knoxville (the latter is the program funder). The mission of the program is to empower parents to become a child's earliest and most influential teacher. The four goals of the Parents as Teachers program are listed below.

1) Increase parent knowledge of early childhood development and improve parenting practices.
2) Provide early detection of developmental delays and health issues.
3) Prevent child abuse and neglect.
4) Increase children's school readiness and school success.

The Parents as Teachers program is designed to serve children from prenatal development through Kindergarten enrollment. Over 240 families have benefited from the Parents as Teachers program to date, though some of those children may have not been old enough to enroll in Knox County Schools at the time of our evaluation. Therefore, we focused our evaluation on those participating students who entered a district elementary school and took the early literacy assessment in Fall 2014. The most reliable measure that can be used to estimate the effects of the Parents as Teachers program would be collected as early as possible in the Kindergarten year. As such, the data from the Fall 2014 early literacy assessment was used in this analysis to estimate differences in Kindergarten readiness among students in families that did not participate.

Findings



The students identified for our evaluation that participated in the Parents as Teachers program had a directionally higher mean normal curve equivalent in the Fall early literacy assessment than students of similar demographics from families that did not participate. Higher normal curve equivalents correspond to higher national percentile rankings for early literacy skills. The magnitude of the difference, however, was not statistically significant.

Group	N	Mean NCE	The probability of the difference in mean NCE between the treatment group occurring by chance
Control 1	8	40	15.9%
Control 2	8	42.8	24.3%
Control 3	8	36.9	10.3%
Treatment	**8**	**56.7**	**N/A**

Recommendations

The Parents as Teachers students for whom we had testing data had directionally positive results. The results of this analysis must be considered in the context of the program. There may be some quality of the families that choose to participate in the Parents as Teachers that predisposes their child to perform better on the benchmark assessments than a student in a family that chooses not to participate. The results of this analysis will not be able to label the Parents as Teachers program as the root cause of any difference in performance.

Additionally, it is very difficult to detect real, statistical differences in mean NCE when analyzing a small sample size. Future analysis of the Parents as Teachers program can utilize data from each prior cohort to increase the number of students participating and the power of any statistical tests used in those analyses.

Source: *University of Tennessee College of Education website*

COMMUNITY SCHOOLS

The Knox County Schools launched the Community School concept at Pond Gap Elementary school in 2011. That project was overseen through a partnership between the school and the College of Education at the University of Tennessee, which provided funding along with other community organizations. Fine arts organizations, churches, and religious organizations have also provided support to the Community Schools program.

In 2012, the concept was expanded to three additional schools: Green Magnet Elementary, Lonsdale Elementary, and Norwood Elementary. SY1314 saw the addition of Christenberry Elementary, Sarah Moore Greene Elementary, and Vine Middle schools. This report did not examine the various services in which the various schools were engaged, but focused on the outcomes exhibited at the seven schools.

Community Schools is a strategy that aligns schools and community resources to provide services that meet the social, physical, cognitive, and economic needs of both students and their families. In particular, it provides enhanced learning opportunities for students and their families via tutoring and mentoring; family engagement activities; health, mental and social services; and early childhood development. This strategy also helps increase cooperation between schools and partners, as well as between teachers and parents.

Community Schools provide services for students that extend beyond the traditional school scope. The program aims to strengthen family and school relations with these targeted, comprehensive services. The community partners provide support to parents and students at the school site to enhance the overall community well-being. The activities available to students and their families are open to the entire school. They include academic and social

programs, as well as access to off-site services within the community. The school-based activities include, but are not limited to, the following:

Student Services	Family Services and Classes
• Academic tutoring • Mentoring • Enrichment classes	• Dinner served nightly • Finance courses • Résumé-writing and interview skills courses • Computer skills courses • GED and ELL (English Language Learners) courses

Findings

While the entire schools were engaged with some Community School activities, the analysis focused on the roughly 725 students who actively participated in the afterschool programs throughout the year. These students were designated as Community School students and their peers as non-Community School students.

The following indicators were used in the analysis:

- Student attendance
- Discipline referrals
- Academic achievement
- Academic growth

For SY1314, the Community School students at Sarah Moore Greene Elementary stand out as being exemplary in terms of overall performance of Community School students. However, this may be due to the fact that, for whatever reason, students that participated in the Community Schools program at Sarah Moore Greene Elementary were predisposed to score better. There is evidence of this "selection bias" because students at Sarah Moore Greene Elementary did not perform as well when considering how their scores changed from one year to the next. When looking at the changes from the previous year where a selection bias cannot have an effect, the Community School students at Christenberry Elementary and Pond Gap Elementary stand out in terms of positive gains in metrics.

Metric	Community School Students Performed Better	Non-Community Schools Students Performed Better
SY1314 Absences	Christenberry Norwood Sarah Moore Greene Overall	
Change in Absences	Sarah Moore Greene	
SY1314 Discipline Referrals		Norwood Pond Gap Overall
Change in Discipline Referrals		
SY1314 RLA Academic Achievement (Proficient or Advanced)	Sarah Moore Greene	
SY1314 Math Academic Achievement (Proficient or Advanced)	Sarah Moore Greene Overall	
SY1314 RLA Academic Achievement (NCEs)	Sarah Moore Greene	
SY1314 Math Academic Achievement (NCEs)	Sarah Moore Greene	Vine
RLA Academic Growth (Proficiency Level)	Christenberry Pond Gap	Lonsdale
Math Academic Growth (Proficiency Level)	Christenberry	Lonsdale Sarah Moore Greene
RLA Academic Growth (NCEs)	Pond Gap	
Math Academic Growth (NCEs)	Christenberry	

Some schools had a smaller proportion of students participating in the Community Schools program than others—a smaller number of participants makes it less likely that we would find any statistically significant differences in metrics. Another way to compare the two groups of students is to apply a grading scale to their average academic growth. Using a scale similar to what the state uses for its report card grades, we get the results seen below.

	Community School Student			
	No		**Yes**	
	RLA	**Math**	**RLA**	**Math**
Christenberry	B	A	A	A
Green	F	F	F	C
Lonsdale	D	B	F	F
Norwood	F	D	C	F
Pond Gap	F	D	C	B
Sarah Moore Greene	F	F	F	F
Vine	C	D	F	F
Total	**F**	**D**	**D**	**D**

Using this representation, the Community School students had better grades in six cells, the same grades in five cells, and worse grades in five cells. Four of the five worse grades came from Lonsdale Elementary School and Vine Middle School. Pond Gap was the one school where the Community School students had better grades in both Reading/Language Arts and Math. Congratulations are also in order at Christenberry Elementary School whose Community School students received an A in each of the subjects.

Recommendations

In its second year of evaluation, the Community Schools program has demonstrated success with regard to intended outcomes. The district should continue the program with its external partners, with regular monitoring to ensure fidelity. Additionally, stakeholders should make a conscious effort to track best practices at the schools that showed the most success in order to replicate them at the other participating schools.

Personalization and Differentiation

In order to achieve excellence for every child, instruction must be personalized and differentiated. This is a priority for KCS and indeed, the state of Tennessee, as it pushes forward its Response to Instruction and Intervention initiative. Personalized and differentiated education is the fulfillment of the first goal from the strategic plan: Focus on the Student—and it requires multilayered instruction and intervention, and screening of students. These efforts also include personal devices and resources to enable student learning, the fourth goal from the 2009 strategic plan.

Focus on the Student

- This goal concerns the role of parents and the community in achieving our vision.
- KCS acknowledges that these stakeholders are important players in supporting the education of all students and that we must build family educational efficacy and coordinate focused district partnerships in order to make that happen.
- Collaborative partnerships with the community will help us achieve the strategic initiatives set out in this plan.

Infrastructure: Enabling Student Learning

- This goal addresses the infrastructure necessary to enable and support the KCS core mission of advancing student academic development and achievement, including effective management of resources that will be essential to achieving our educational goals.
- KCS will leverage its assets and resources to operate efficiently, effectively, and professionally.
- KCS will ensure a safe, healthy, inviting learning environment for all students, and base our actions on sound data and accurate, relevant information.

Source: Knox County Schools Strategic Plan (2009-2014)

These two goals have been a catalyst for KCS to commit greater resources towards implementing intervention programs and exploring new ways to impact student outcomes. This section is dedicated to three such programs: two literacy intervention programs utilizing Voyager Passport in the elementary grades and the School Technology Challenge (STC) pilot in 11 schools district-wide.

Investment Analysis

The intervention programs were budgeted to support both personnel expenditures and materials. Some considerations to note as part of this investment analysis include the following:

- All costs presented include personnel services and benefits.
- The Early Literacy Initiative costs reflect the portion funded by the County Mayor and do not include $90,317 in carryover funds for FY15.
- Student counts encompass those who benefitted from the additional supports or were part of the program evaluation.
  - The student count for early literacy initiative includes all students receiving intervention services, though their materials may have been purchased prior to FY14.
  - The AERS line item funded 20 instructional assistant positions. Therefore, the AERS student count includes only those students provided intervention services by the instructional assistants who were hired through this funding.

Initiative	FY14 Expenditures		Total FY14 Expenditures	# of Students	Cost Per Student
	Early Literacy	General Purpose			
Early Literacy Initiative	$2,870,000		$2,779,629	2,835	$980
Additional Elementary Reading Support	$324,174	$15,612	$339,786	755	$450
School Technology Challenge		$2,781,144	$2,781,144	16,166	$172
PERSONALIZATION & DIFFERENTIATION	**$3,194,174**	**$2,796,756**	**$5,900,559**	**19,756**	**$299**

EARLY LITERACY SUPPORT

Voyager Passport is the reading intervention program provided through district resources. All 49 elementary schools participated in this intervention. Students receiving the intervention support participated in an additional 30 minutes of reading instruction a day.

In our previous EROI report, the SY1213 literacy intervention results were disappointing overall. One group that showed some positive intervention results came from the population for whom the intervention was originally intended, that was students in grades 1-5 who scored between the 11th and the 24th percentiles on the AIMSweb CBM reading (R-CBM) data. For SY1314, students at all 49 elementary schools were chosen for the intervention based on multiple data points that included the Spring 2013 TCAP percentile and the Fall 2014 administration of the R-CBM. Students in grades 1-5 who scored between the 11th and the 24th percentiles were still intended to be the intervention participants in conjunction with the forthcoming Response to Instruction and Intervention (RTI^2) initiative. The intervention consisted of students receiving an additional 30 minutes of reading instruction with Voyager Passport. Classroom teachers and/or instructional assistants were to provide the instruction. Voyager Passport is an intervention program intended to assist students with word study, fluency, comprehension, vocabulary, writing, listening and speaking.

Our intent was to test Voyager student growth as measured by the difference between the observed scores and the predicted scores over a one-year period. Our previous report did this using three separate measures: SAT-10 scaled scores for grades 1-2, TCAP Achievement scaled scores in grade 3, and TCAP NCEs in grades 4-5. For this report, we did not combine any grades and considered each grade separately.

Superintendent's Report

In a January 2015 report to the County Mayor from the Superintendent, the Early Literacy initiative was evaluated based on two years' of data (whereas our evaluation is based on one year). The data presented was an aggregate of the district and the schools that received early literacy support. That report noted positive gains were made possible by the early literacy grant provided by the County Mayor. Its findings include the following:

- In grades K – 2, the district overall improved in terms of mean scaled score and median national percentile rank on the SAT-10 from 2013 to 2014.
- In SY1213, the district made positive gains in reading and language in grades 1 – 2. However, in SY1314, grades 1 – 2 did not show positive growth in reading but language remained positive.
- In grades 3 – 5, student achievement was up in SY1213 but saw decreases in SY1314.
- SY1213 student growth was positive in grades 3 – 4, though in SY1314 only grade 3 student growth remained positive.

- The schools directly receiving early literacy supplemental funding in both SY1213 and SY1314 performed better than the schools that did not in terms of student growth.

The full report by the Superintendent is available in Appendix D.1.

Findings

Our findings are based on one year (SY1314) of data that we were able to collect for the students directly impacted by the early literacy initiative through the Voyager Passport reading intervention. While the Voyager Passport program was intended to be used with students in the 11th to the 24th percentile, we found that many students from this percentile band were not in the intervention. When Voyager and non-Voyager students were tested against one another as a whole, the growth in TCAP and SAT-10 scores was mixed.

In an effort to remove as much potential bias as possible, an analysis was conducted between demographic-equivalent students. With a very large sample of equivalent students, the non-Voyager students outgained the Voyager students significantly in four of five grades as measured by TCAP and SAT-10 test scores. We have noted that the selection process for placing students in intervention is not consistent across the district, as various subjective methods are being used along with the screeners. It is possible that this adds to the mixed results present in our findings. The results of our analysis indicated that the Voyager Passport reading intervention is not having a positive effect on the district's Reading/Language Arts scores on state assessments.

Recommendations

Similar to our previous evaluation of early literacy support, there were many students enrolled in intervention for reasons we could not readily discern. It is important for school administrators to document rationale as to why students are placed in Voyager intervention. Documenting such decisions will most likely be less of an issue as the district continues to fulfill the requirements as part of RTI2.

ADDITIONAL ELEMENTARY READING SUPPORT

Twenty schools were provided with an instructional assistant to improve Reading/Language Arts scores and to help facilitate a reading intervention with designated students. While the program was initially designed for grades three to five, we found over 100 students in grades one and two who also received assistant-led intervention. This analysis is a smaller version of the Early Literacy Report with a focus on the students supported by interventionists.

The interventionists maintained logs of the students assigned to them. Some schools incorporated dynamic scheduling and altered the person providing the intervention during the year between a certified teacher and the AERS interventionist. From the intervention logs, we found 755 students who were at one point assigned to the AERS interventionist. We linked state assessment scores to the predicted scores for the 604 students who took the Spring 2014 state assessment and had a sufficient test history to generate a predicted score. As the state assessments differ slightly from grade to grade, we examined each grade separately.

We defined growth to be the difference between the observed score and the predicted score and then compared the growth between the AERS students and other students. We first compared the students as a whole for the 20 schools. We then generated control groups based upon demographic features and predicted scores and evaluated the AERS students using matched-pair designs.

Findings

The AERS intervention exhibited mixed results. Overall, in four of the five grades, the mean student growth for the AERS students is smaller than the mean growth of the rest of the students in the particular grades. Yet, when broken down by school, there are grades where the AERS students performed better than the other students in their grade based upon the growth means. We did not see any positive results for the AERS student growth when matching them with students who had similar predicted scores. This may be because once they are matched, the control student may also have been in an intervention; and if this intervention was in an AERS school, then the control student would probably have received his or her intervention from a certified teacher instead of from an educational aide.

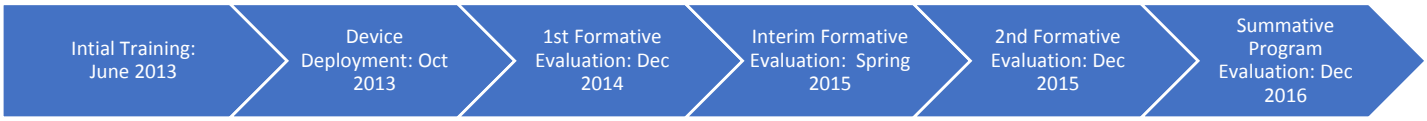
Recommendations

Program leaders should continue monitoring and reporting on key metrics as part of any budgetary compliance. The district should also consider ways to better track complementary students to use as control groups and determine if those students were in also in intervention programs for evaluation purposes. A further study would require details on all students who receive intervention, the nature of the intervention, and the certification status of the intervention provider.

SCHOOL TECHNOLOGY CHALLENGE

In SY1314, Knox County Schools made a significant investment in computing devices and associated professional development to fully integrate technology within a subset of 11 KCS schools. The resulting School Technology Challenge (STC) created a 1:1 student to device ratio with the main objective of increasing teacher effectiveness to drive increases in student outcome data. The STC theory of action highlights three areas of focus for meeting this main objective. The foci are listed below.

1) Increase the individualization and differentiation of student-centered instruction.
2) Increase student affinity, motivation, and engagement in the classroom.
3) Increase the effectiveness of teaching through both 1) and 2) while integrating technology-based education aids.

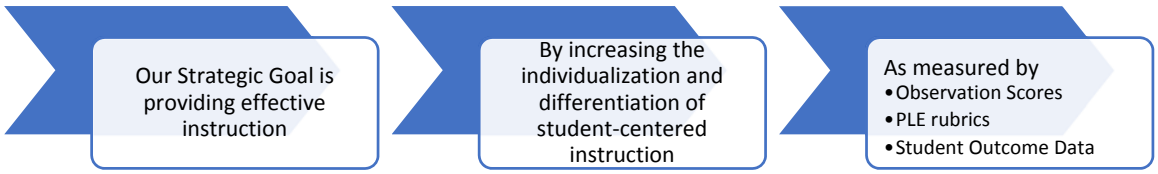


Findings

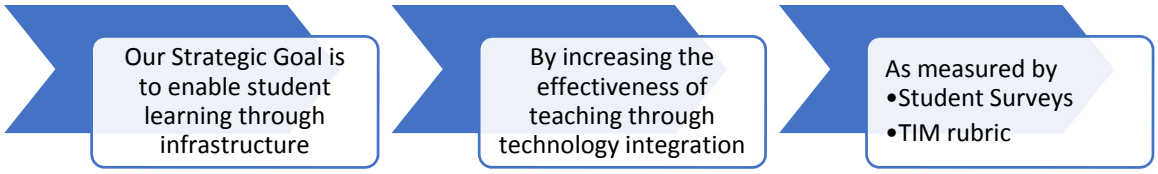
The SY1314 formative analysis indicates that the STC is very clearly a work in progress. Data collected through focus groups, surveys, and classroom observations indicate that the depth of technology integration is likely not yet deep enough to move the needle on many key student outcome indicators. Results relating to the STC theory of action are highlighted below.

- **Increasing Student Engagement, Motivation, and Affinity:** Common themes from teacher focus groups indicated that students were more engaged in classwork when the technology was used in the classroom. Student responses to pre and post-deployment surveys corroborate this perception. TRIPOD data provides further evidence of student engagement as the Captivates domain was directionally higher for STC schools (compared to control schools).

- **Attendance:** There was no statistical difference in the change in attendance rates between STC and control schools. There is little evidence to indicate that the STC had any impact on student attendance rates.



- **Increase the individualization and differentiation of student-centered instruction:** Rubrics regarding Technology Integration (TIM) and Personalized Learning (PLE) were scored from a pool of 112 randomly selected teachers. The majority of teachers scored in the earliest stage of technology integration. Teachers scored better on the PLE matrix, but almost all data indicated that teachers were operating in an "emerging" state of personalizing the learning environment. The results echo common themes from the teacher focus groups. Respondents felt that this year was more of an experimental year to determine what processes worked in the classroom and what processes did not.



- **Increase the effectiveness of teaching through technology integration:** Since the depth of technology integration and personalization does not appear to be sufficient to fundamentally change the classroom experience, it is probably too early in the process to expect large changes in student outcome data. TVAAS data provides some directional evidence of school-wide increases, but the effects are not statistically different than the control group. There is no evidence yet of any systematic closure of performance gaps and results regarding achievement data are similarly mixed. Focus group respondents felt they were better prepared to enter year two of the STC with strategies to deepen personalization and help significantly impact student outcomes.

Recommendations

It is important to note that the School Technology Challenge has always been viewed as a multi-year project. This analysis can serve as a formative sign-post for the initiative, but it is too early in the life cycle of the project to determine its true worth to the KCS.

Based on the timeline for program implementation and review, the REA team makes the following recommendations in regard to the School Technology Challenge.

- KCS should continue to focus on continuing professional development, training, and promoting teacher-to-teacher collaboration to maximize the benefits of the STC.
- Available formative data and mid-year rubric scores should form the backbone of an interim formative analysis on STC implementation. Note, however, that formative data will be universally available only for elementary schools.
- Program leaders and evaluators should present the findings of the SY1314 formative program evaluation to the teachers at participating STC schools and begin the SY1415 cycle of data collection through teacher focus groups in Spring 2015.

Teacher Support

Instructional coaches and lead teachers are just two ways in which the KCS aims to pursue its goal of effective educators from the 2009 strategic plan. The goal is to have an outstanding teacher in each classroom, highly effective leaders in each school, high quality instruction and support for teachers, and a culture of instructional excellence in the district. This theme was carried forward in the newly adopted strategic plan (2014-2019), which also calls for building the KCS community of learners by increasing instructional capacity of educators. In the SY1314, there were 142 instructional coaches and 226 lead teachers working in schools across the district.

Effective Educators

- This goal is focused on ensuring there is an outstanding teacher in each of our classrooms and a highly effective leader at the helm of each of our schools.
- Appropriate support will be provided to the teachers and principals in order for them to effectively do their jobs.

Source: Knox County Schools Strategic Plan (2009-2014)

We have broadly classified this section as "Teacher Support" since that is the aim of instructional coaches and lead teachers. We evaluated coaches via their work in individual learning cycle (ILC) support and qualitative feedback surrounding professional learning communities (PLCs). We also tabulated observations conducted by lead teachers as a proxy of the support they provide building administrators.

The learning cycles that instructional coaches facilitate, and which are the basis of our evaluation, are outlined below. We also conducted focus groups to gather qualitative data to layer on quantitative data to more holistically analyze the district's coaching model.

Learning Cycles at a Glance		
	PLC Cycle (PLCC)	**Individual Learning Cycle (ILC)**
Content	Literacy Numeracy Pedagogical content knowledge	TEAM Instructional Indicators
Participants	All teachers (staff development days) Teacher cohort (weekly PLC support)	Individual teachers
Teacher selection	Grade level teams (by department in secondary) are assigned to PLC cohorts by the principal.	Teachers are assigned by principal based on TEAM data, with input from coach, lead teacher, and other observers.
Start date	August	August
Length	6-9 weeks	6 weeks
Number of cycles	4 per coach per school year*	5 per school year*
Structure of cycle	Learning experience and planning (all) **Classroom support (cohort)*** **Debrief (cohort)*** **Weekly collaborative reflection, learning experience, and planning (cohort)*** Evaluation – DuFour's Audit Rubric	Observation Develop coaching plan **Learning experience and planning*** **Classroom support*** **Debrief*** Observation and evaluation
Facilitator	Instructional coach	Instructional coach
Caseload	A coach should have no more than 10 – 15 teachers in an individual PLC. Coaches will be responsible for up to 4 PLCs per week, based on school needs and resulting prioritization.	No more than 4 teachers per learning cycle per coach.

Source: Continuing the Journey: Coaching & Learning Cycles Handbook. Knox County Schools.

The vast majority of coaches specialize in either literacy or numeracy, with two system-wide coaches to support science and social studies. Coaches facilitate PLCs and ILCs, provide support to school administrators and teachers, and attend monthly Coaches Network professional development workshops.

Updates from EROI 2014

Based on evaluation within the Curriculum and Instruction department, as well as reflection on the results of the previous EROI (published April 2014), there were modifications made to the coaching model during SY1314 and for SY1415:

- Based on feedback from school administrators, teachers, and coaches, principals now have oversight of the coaches in their building (as of SY1415).

- A book study of *The Art of Coaching* is part of the Coaches Network for SY1415 to address the needs of coaches with regard to training for instructional coaching techniques and adult learning.
- TEAM evaluation rubric training is now a component of the monthly Coaches Network meetings in SY1415 to help deepen coach knowledge of the rubric.
- The ILC intake form was modified in SY1314 in order to capture more information about the participants and the duration of the cycles.

Investment Analysis

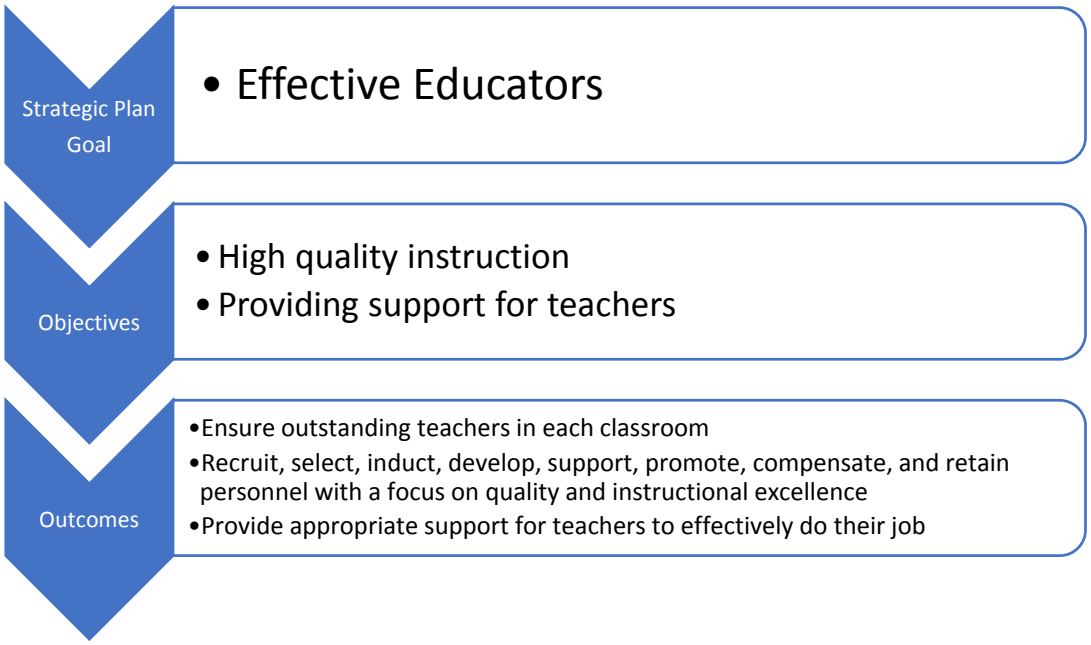
The cost for teacher support is represented as a "per teacher" expenditure since the staffing ratios are typically driven by the number of teachers or certified staff at the location versus student counts. In the case of coaches, they were typically allocated per school and program, which is why the range of coach to teacher ratio spanned from 1:9 to 1:200. The number of teachers supported by lead teachers represents all teachers in TEAM schools only. Instructional coaching supports teachers in all schools in the district.

The column marked as Federal and/or Grant includes funds from the federal and state governments, such as: Title I, Title II, and Title III, the Innovation Acceleration Fund, and Race to the Top dollars. All costs presented include personnel services and benefits.

Initiative	FY14 Expenditures			Total FY14 Expenditures	# of Teachers	Cost per Teacher
	Federal and/or Grant	Early Literacy	General Purpose			
Lead Teachers	$291,050		$379,790	$670,840	3,459	$194
Instructional Coaches	$6,598,884	$1,735,378	$1,470,000	$9,804,262	4,447	$2,205
TEACHER SUPPORT	$6,889,934	**$1,735,378**	**$1,849,790**	**$10,475,101**	**4,447**	**$2,356**

INDIVIDUAL LEARNING CYCLES

The use of instructional coaches and learning cycles is part of the strategy to meet the Effective Educators goal from the district's five year strategic plan.



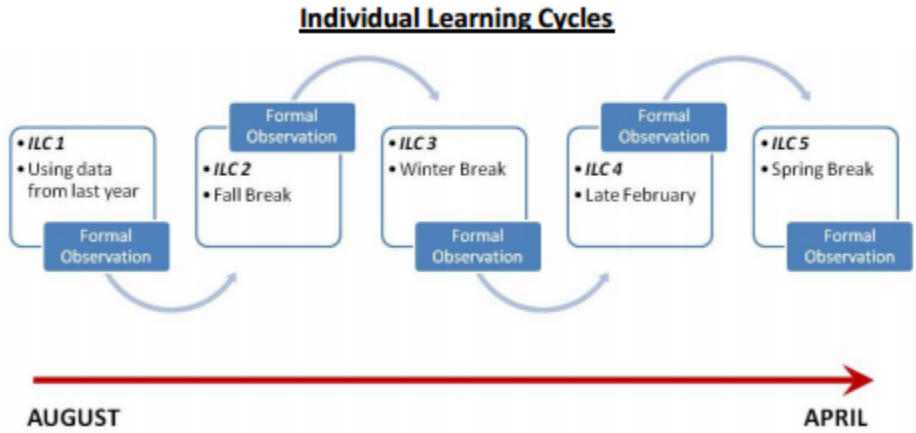
Adaptation of Goal 2 from the KCS Five Year Strategic Plan (2009-2015)

Coaches work one-on-one with certain teachers via individual learning cycles (ILCs). Coaches are asked to spend at least a quarter of their work time on preparing for and conducting ILCs. ILCs are one part of the overall learning cycle for teachers. According to the KCS Coaching Handbook (*Continuing the Journey: Coaching & Learning Cycles Handbook,* August 2014), the key outcomes of the learning cycles are

- Instructional coaches use a consistent, research-based approach to professional development (PD) that includes structured follow-up and opportunities of teacher collaboration.
- Teachers receive training and support in high-impact instructional strategies aligned to the Common Core State Standards transition and the TEAM instructional indicators and implement them with fidelity and understanding.
- All KCS stakeholders understand the role of the instructional coach.
- Coaches receive intensive training in content, PD delivery models, TEAM, and cognitive coaching competences.

Teachers are placed in ILCs based on low scores (Level of Effectiveness, classroom observation, individual growth), principal recommendation, or teacher self-selection. ILCs are based on a six-week cycle that includes a formative classroom visit, a coaching plan, implementation of the coaching plan, and reflection and feedback. ILCs are opportunities for

the coach and teacher to craft a personalized plan to help improve the teacher's instructional delivery and student outcomes. The coach and teacher collaboratively determine the focus of the ILC and coaching plan. ILCs typically include: modeling, co-teaching, co-planning, observing, and providing feedback. ILCs focus on refinement areas and goals for the teacher, including classroom evidence of fulfillment of those goals. ILCs are to end with "next steps" for both the coach and teacher. A typical ILC is based on the following schedule (see the figure below). In short, instructional coaches are expected to provide teachers with high-quality learning experiences and intensive classroom support.



Source: Continuing the Journey: Coaching & Learning Cycles Handbook. Knox County Schools.

Findings

We wanted to estimate if observation scores and student outcomes (as measured by TVAAS gains) improved as a result of teacher participation in an ILC. Before reviewing the findings of the ILC analysis, however, it is important to consider the following:

- The majority of teachers enrolled in an ILC earned scores below expectations. Therefore, teachers in an ILC may be predisposed to perform poorly or be rated poorly in the classroom, regardless of intervention.
- Classroom observation requirements changed mid-year in SY1314, which may have had an inadvertent but negative impact on observation scores.

There were approximately 463 teachers who participated in an ILC in SY1314, 327 of whom had prior evaluation data to use as a comparison point. We compared ILC teacher evaluation data from SY1213 (prior to ILC treatment) to SY1314 (after completion of ILC participation). We focused on ILC teachers who fell in the target of range for intended ILC participation (that is, those with Level of Effectiveness or individual growth scores of Level 1 or 2). We also looked at those critical target ILC teachers, who earned scores of Level 1 only. We also compared ILC teachers to similarly situated peers in a control group. The control group

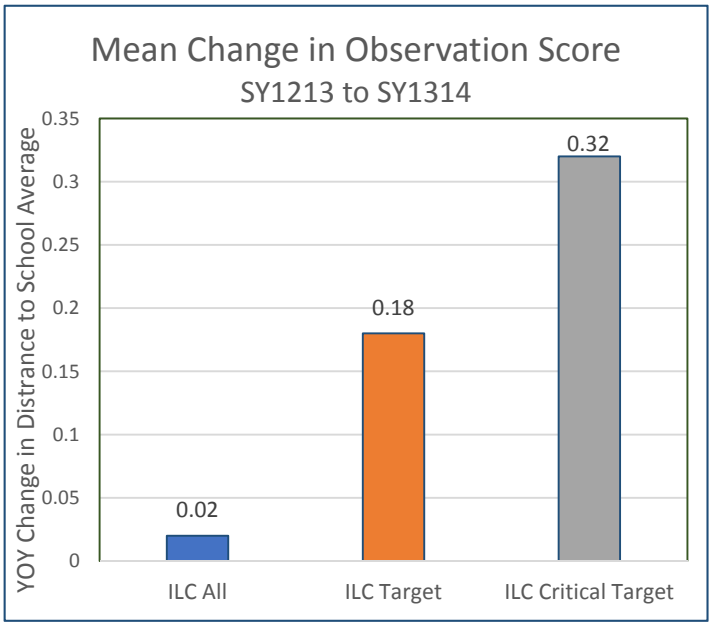
teachers had similar years of service and prior evaluation data as the ILC teachers. The following table summarizes the results of our analysis.

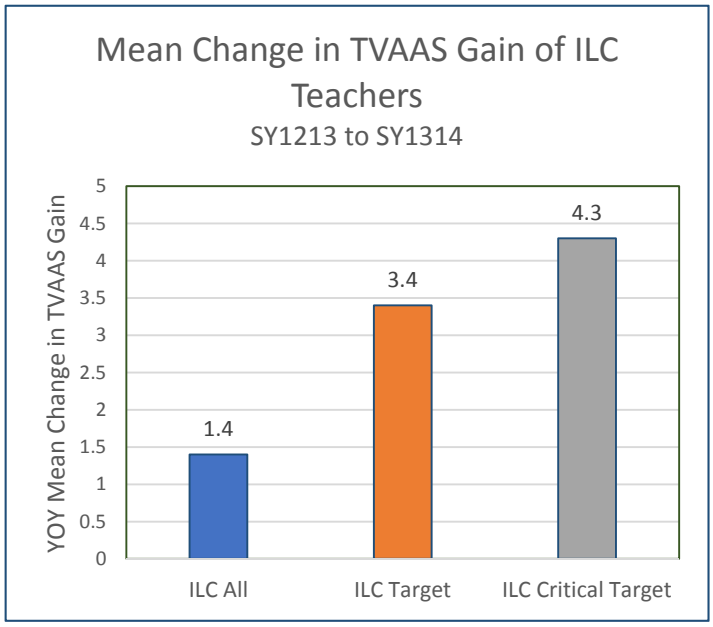
Unit of Analysis	Classroom Observation Scores	TVAAS Gains
Intended ILC Target	Increased*	Increased* but still below growth standard
All ILC Teachers	Below school average	Increased* but still below growth standard
ILC & Control Group	Control group outperformed ILC*	Control group outperformed ILC*
ILC & Non-ILC Apprentice	Increased but no discernible difference from Non-ILC Apprentice Teachers	Increased but no discernible difference from Non-ILC Apprentice Teachers

*denotes statistical significance

The results of this analysis suggest that ILCs have positive impacts for those teachers who qualify for ILCs based on low scores. Based on the quantitative metrics we used, the strongest indicator of improvement was found among those ILC teachers who qualified for ILCs. There was no evidence to suggest that apprentice teachers who were enrolled in an ILC performed any better than their non-ILC apprentice peers.

- Considering only those intended target ILC teachers who qualified for ILC participation based on their scores:
    - The mean classroom observation score increased, although these ILC teachers still earned scores below their school average.
    - They also increased their mean TVAAS gain, though it was still below the growth standard (a mean TVAAS gain of 0).
    - When looking at the ILC teachers who qualify as critical targets (starting observation score less than 2.55 and Level 1 individual growth score), they showed even more significant growth after treatment.





  - Their mean classroom observation score was below the average classroom observation score in their schools.
  - They increased their mean TVAAS gain from a mean of -3.5 in SY1213 to a mean of -2.1 in SY1314. The difference between the means was statistically significant. While it is an increase, their TVAAS gains are still below zero.
  - Half of the ILC teachers earned the same individual growth score in SY1314 as in SY1213 while 27% improved their individual growth score in SY1314.
- Considering ILC teachers in comparison to similarly situated peers:
  - The control group improved their mean classroom observation score at a higher rate than the ILC teachers and the difference was statistically significant.
  - ILC teachers had a mean TVAAS gain of -1.46 while the control group teachers had a mean TVAAS gain of -0.15. The difference between the means was statistically significant. While both groups had negative gains, the control group mean was not as far below the growth standard as that of the ILC teachers.
- ILC and non-ILC apprentice teachers' outcome data was almost identical.

In addition to quantitative data points, we also used three qualitative sources of perception data: the TELL survey, a district-wide survey of KCS staff, and focus groups with instructional coaches.

- The questions in the TELL survey were not specific to ILCs, but inferences were made based on the responses to the survey. Teachers that participated in the TELL survey reported that: they had less time to collaborate but more access to professional

support personnel; professional development did not enhance their ability to help improve student learning and it did not deepen teacher content knowledge; they need less professional development; and fewer new teachers reported receiving support in 2013 than in 2011 (the first administration of the TELL survey).

- The questions in the district-wide survey were specific to instructional coaches, but not ILCs. Inferences made from the survey results are based on the respondents' perception of coaching. The district-wide survey indicates respondents felt the most impact from instructional coaches in the areas of supporting collaboration and providing access to and encouraging the use of different and new resources. Teachers who were surveyed responded less favorably about the impact instructional coaches had on learning the evaluation rubric or changing knowledge of instructional practices.
- The feedback from focus groups included in this report was ILC-specific. Coaches who participated in focus groups reported that their efforts are most effective in schools wherein the administrators are very supportive of instructional coaches and learning cycles. The coaching model outcomes call for coaches to receive intensive training in content areas, professional delivery models, TEAM, and cognitive coaching competencies. The feedback from coaches suggests that these outcomes have not been fully reached.

The qualitative data indicated that while teachers perceive instructional coaches to be helpful, the impact of coaches was not as clear on outcomes related to the learning of the evaluation rubric or changing instructional practices. These findings may be due, in part, to changes in observation requirements that occurred mid-year in SY1314: all professionally licensed teachers had unannounced observations in Fall 2013, but were then allowed to choose announced observations in Spring 2014.

Recommendations

The evidence points to the conclusion that ILCs have positive impacts for those teachers who earn low scores and ILCs have a negative impact for teachers at or above expectations. Therefore, the district may wish to narrow the focus of ILCs to low-scoring teachers and provide different educational opportunities for teachers at or above expectations. If the district chooses to continue monitoring and investing in the coaching model at the central level, the recommendations moving forward are as follows:

- The district should consider limiting ILCs to those teachers with a Level 1 or 2 score, since that group of ILC teachers showed significant improvement.
- Similarly, the district should consider limiting the number of effective teachers in ILCs. Almost 20% of teachers who were selected to participate (not self-selected) in an ILC in SY1314 were professionally licensed, considered "effective" (they had a Level of Effectiveness score of 3, 4 or 5), and had classroom observations score of 3.0 or higher in SY1213. It is unclear why these effective teachers were enrolled

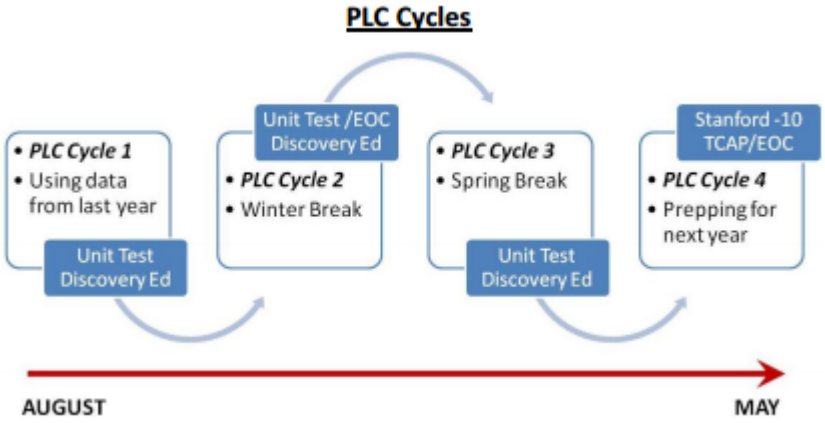
in an ILC—even if they self-selected to enroll—they should have been directed to different resources rather than an ILC, such as PLC leadership or the mentoring program. ILCs did not have the intended effect on this group of teachers.

- The district may wish to bolster its current mentoring program (funded by the Great Schools Partnership) by connecting new teachers with highly effective teachers in their buildings. This would free up coaches' time for professionally licensed teachers who have earned scores below expectations.
- The district should continue to build upon its current (SY1415) practice of providing instructional coaches with additional and consistent TEAM rubric training opportunities, adult learning training, and content-area training.
- The district may wish to develop best practices surrounding the ILC process, including (but not limited to)
    - A meeting between the principal and coach to discuss the areas of refinement (from the evaluation) for each teacher in an ILC so that the coach can plan to address those areas.
    - An initial meeting between the principal, teacher, and coach should be scheduled prior to or during the first meeting of the ILC to establish a positive and productive perception about the process. This would also provide an opportunity for the teacher to ask questions about the process if s/he is unfamiliar with it.
    - Coaches should, whenever possible, work with teachers within their content specialty.
- Since the referral process is based on multiple data points and several decision-makers, the data collection form should be completed such that there is clear delineation why a teacher is selected for an ILC.
- In addition to reinforcing the referral collection form, data on how often teachers and coaches are meeting should be uniform and collected centrally.

PROFESSIONAL LEARNING CYCLES

Instructional coaches provide school-based, job-embedded professional development for teachers in order to raise the quality of teaching and learning across a school. The work of coaches is also intended to build collective leadership to improve outcomes for students. Instructional coaches typically model lessons, provide and interpret data with principals and faculty, facilitate PLCs, and help with the intervention process (screening and working with students). Coaches are expected to spend 50% of their time preparing for, facilitating, and conducting PLCs.

PLCs are an opportunity for coaches to help teachers plan, collaborate, and use data to inform instructional decisions and plans. The goal is to raise the quality of teaching and learning across a school and build up teacher leadership skills. Typically in PLCs, coaches model lessons, provide and interpret data with PLC members, and they may also participate in intervention decisions. Typically, school leaders, the coach, and teachers create a nine-week instructional plan, implement the plan, analyze the results, and adjust instruction based on those results.



Source: Continuing the Journey: Coaching & Learning Cycles Handbook. Knox County Schools.

As part of the Teacher Support initiative and the Effective Educators strategic goal, instructional coaches are asked to participate and sometimes lead meetings of professional learning communities (PLCs). Unlike ILCs, PLCs are supposed to include all teachers in a school. The strategic plan frames PLCs as an opportunity for teachers to collaborate with one another using data and sharing best practices while also leveraging expertise within buildings. Coaches are meant to assist in content knowledge and instructional delivery.

Updates from EROI 2014

One major component of PLCs is SMART goals, which guide the cycle planning, implementation, and results. The evaluation of PLCs in the previous EROI looked at the rate of SMART goal attainment and its effect on the mean TVAAS growth index. The results

suggested that schools that achieved a higher percentage of their SMART goals in PLCs also had a higher TVAAS growth index, but the results were not statistically significant. The evaluation also emphasized that the quality and content of SMART goals was difficult to assess at a central level. Therefore, the process of collecting SMART goals centrally was discontinued in SY1415. Since SMART goals are no longer centrally collected and monitored, this evaluation did not focus on SMART goals or the impact of their attainment on student outcome data.

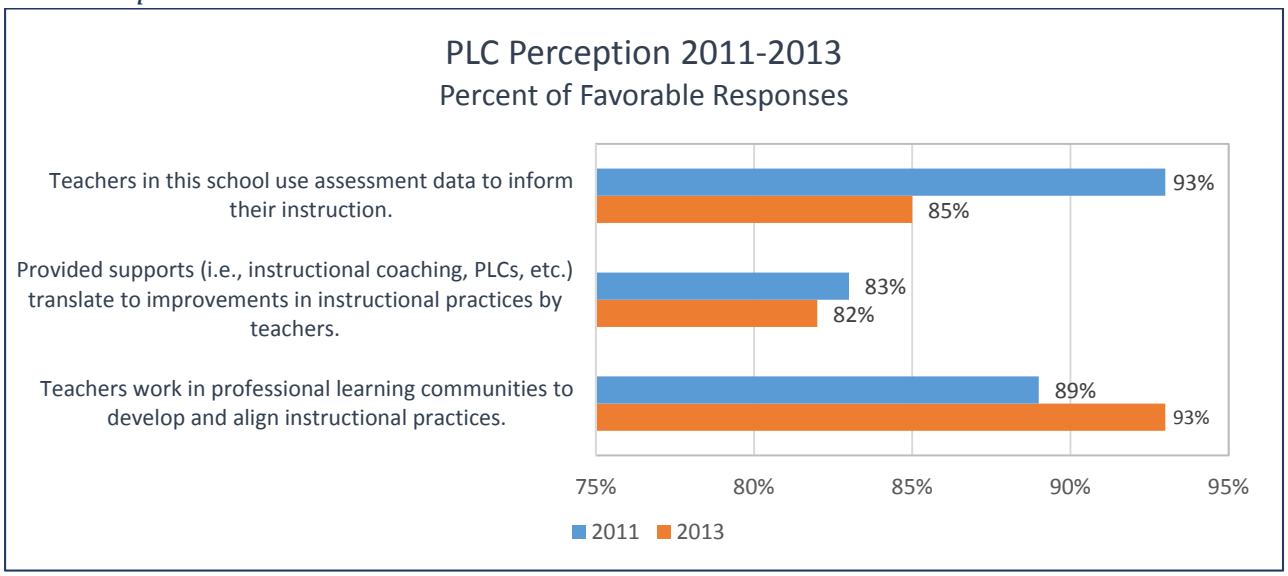
Findings

The focus of this evaluation was qualitative data regarding the effect of PLCs. We triangulated data from three different sources to inform our findings. Each source of qualitative data is outlined in the subsections below.

TELL Survey

The TELL (Teaching, Empowering, Leading and Learning) Tennessee Survey was launched in 2011 and is administered every two years. According to its website, the survey provides educators with data to facilitate school improvement. It includes questions from a range of topics, including teacher leadership, use of time, professional development, and instructional practices and support. The results from the TELL survey are especially helpful because the first administration was prior to the revamp of the coaching model in the district. In 2011, almost 73% of all KCS teachers responded to the survey, while in 2013 only 44% responded. We captured favorable responses by combining the amount of "agree" and "strongly agree" responses to the survey questions.

PLC Perceptions

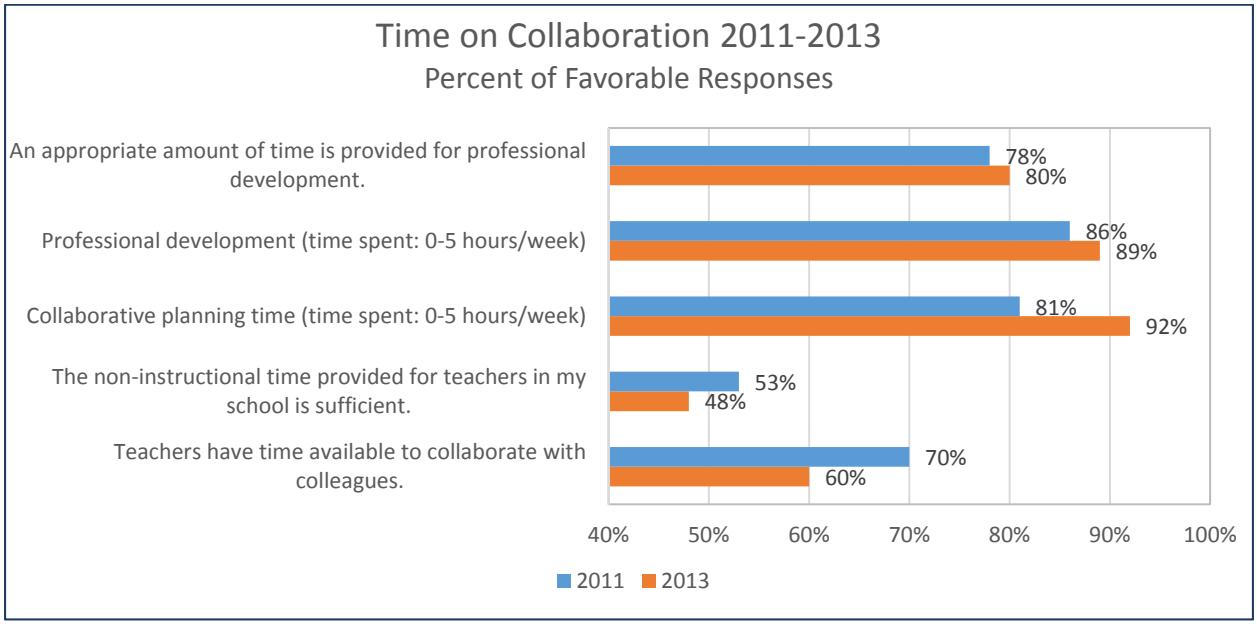


When looking at questions related to PLCs, there was a decrease in the perception around use of data and impact of PLC on instructional practice from 2011 to 2013. However, the

percent of favorable responses to the statement regarding PLCs serving their intended purpose increased by 4% since the revamp of the coaching model. These results suggest that educators perceive that PLCs are used, and used properly. However, educators do not rate the outcomes and impact of PLCs as highly. Nonetheless, there was an overwhelming majority (over 80%) of teachers surveyed responding favorably with regard to PLCs.

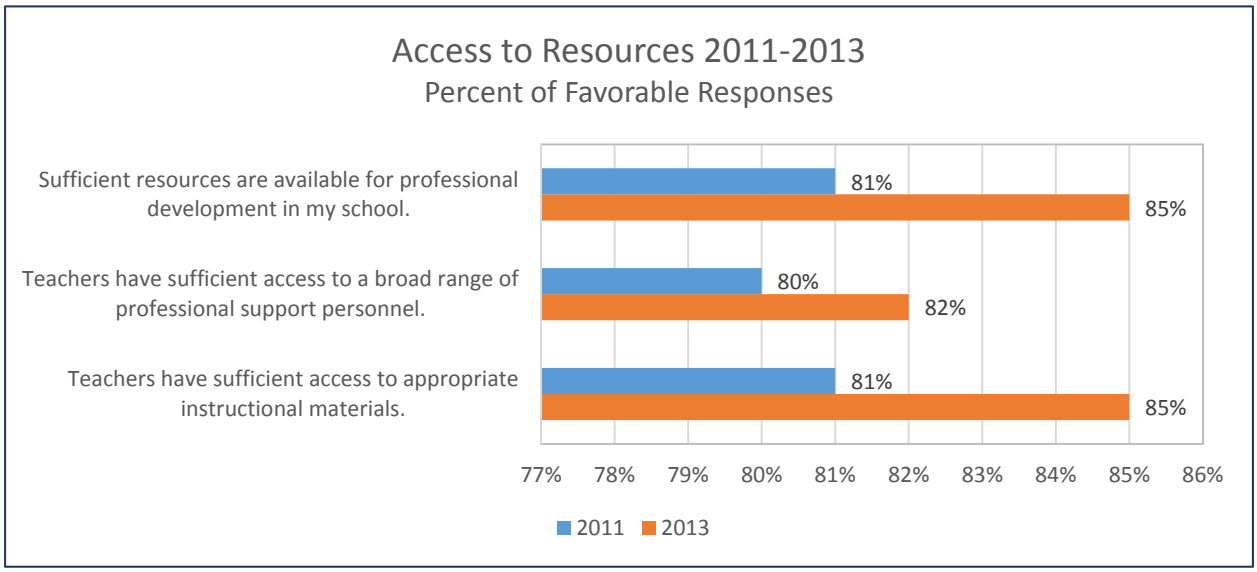
Time Spent on Collaboration and Professional Development

The questions including actual time spent were tallied based on the percent of educators that reported 0-5 hours per week on the specified activities. There was an increase in the amount of time provided for professional development and collaborative planning time. However, educators reported having less non-instructional time as well as collaborative time with colleagues. These mixed results may indicate that while teachers did have time to collaborative planning time via PLCs, they did not have adequate time to collaborate outside of PLCs with colleagues (perhaps planning periods).

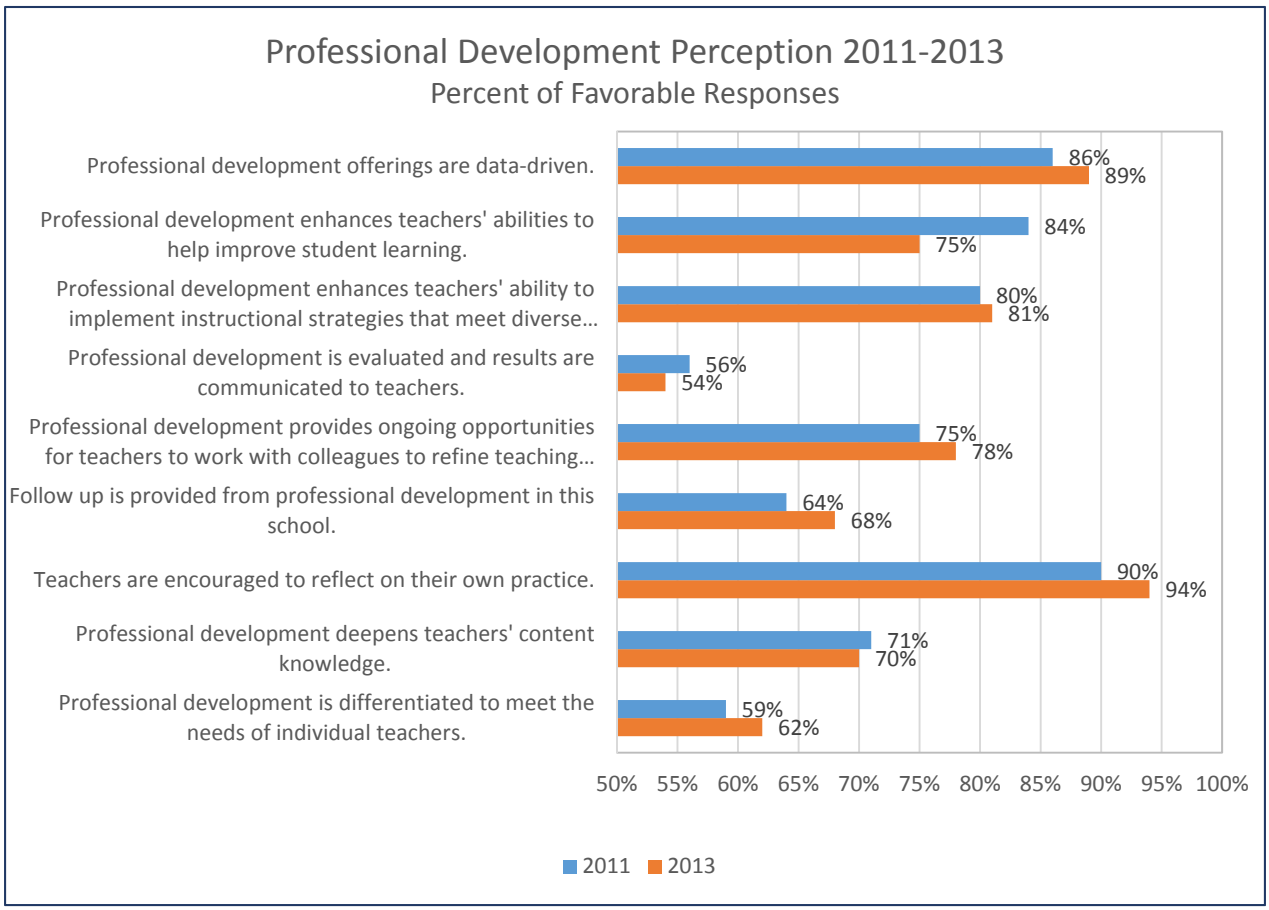


Access to Resources

There has clearly been an increase in favorable educator perception of access to resources for professional development, personnel, and materials. While responses were highly favorable in 2011, they increased by an average of almost three percentage points.



Professional Development



Educator responses regarding professional development were varied. There seems to be a perception that while learning cycles are useful, their impact does not translate into

improved student outcomes. This perception was also recorded in the district survey results (next section). Overall, there was a positive perception about the use of data and an increase in favorable perception regarding collaboration and reflection.

District Survey

In May 2014, the REA administered a district-wide survey with questions intended to gather perception data about several different departments and district initiatives, including instructional coaches. The survey was sent to a random selection of school administrators, classroom teachers, instructional assistants, instructional coaches, and other certified personnel. The questions were related to instructional coaches, so inferences about PLCs were based on coaching perceptions. There were seven instructional coaches that responded to the survey, but given the small n-count, their responses are not included in the table below.

In order to analyze the responses, we combined "agree" and "strongly agree" responses as favorable responses. We tallied all administrator and classroom teacher responses separately. The table below shows the percent of favorable responses by administrators (114 surveyed) and teachers of all grade levels (an average of 341 surveyed).

Question	% of Favorable Responses	
	Administrator Respondents (n = 114)	Teacher Respondents (n = 341)*
The instructional coaches led teachers to think about an aspect of their teaching in a new way**.**	81%	53%
The instructional coaches encouraged teachers to pay closer attention **to particular things that were being taught.**	80%	52%
The instructional coaches led teachers to seek out additional information or other resources**.**	82%	53%
The instructional coaches encouraged collegiality and collaboration **among teachers.**	84%	58%
The instructional coaches led teachers to question their beliefs and assumptions about which teaching methods **work best with students.**	76%	42%
As a result of interactions with instructional coaches, have there been changes in classroom management practices**?**	54%	21%
The instructional coaches led teachers to modify or improve the lesson planning process**.**	80%	50%
As a result of interactions with instructional coaches, have there been changes in teacher knowledge and understanding of instructional practices**?**	67%	36%

Question	% of Favorable Responses	
	Administrator Respondents (n = 114)	Teacher Respondents (n = 341)*
As a result of interactions with instructional coaches, have there been changes in teacher knowledge and understanding of the evaluation rubric**?**	59%	36%
As a result of interactions with instructional coaches, have there been changes in the use of student data in instructional planning**?**	71%	40%
As a result of interactions with instructional coaches, have there been changes in the use of differentiated instructional strategies**?**	66%	40%

*The n-count for teacher respondents varied by question (range of 338-343) with a mean count of 341.

The results of the district-wide survey administered at the conclusion of SY1314 suggest that principals perceive the impact of instructional coaches more favorably than classroom teachers. Administrator respondents rated collaboration most favorably (84%) while classroom management practices only garnered 54% favorable responses. The same was true for teacher respondents, though classroom management practices only received 21% favorable ratings. The positive responses regarding collegiality and collaboration can be a good indicator of educator perception of PLCs, given that PLCs are one of the key vehicles of collaboration in schools.

Focus Groups

In April and May 2014, the REA hosted a series of focus groups with instructional coaches from across the district. More than 30 coaches participated and each specialty (elementary, math, literacy, match, GT, etc.) was represented. They were asked six questions related to the coaching model, their needs, and the perception regarding the impact of PLCs. These six questions covered different topics, but the ones pertaining to PLCs included the following:

- What kind of training would better prepare you for your job?
- What are you doing differently from last year, from other coaches?
- How are you spending your time? Describe a typical week.
- Describe the difference between ILCs and PLCs both in how they are implemented and their impact on teachers.

Training for Coaches

Coaches requested more time to collaborate with each other.

Coaches requested more training in content areas, adult learning, and the TEAM evaluation rubric.

Coaches requested the ability to visit other schools to see how their peers are doing things.

Differences from Previous Years

More coaches reported planning sessions with principals at the beginning of each week.

Coaches reported using surveys to collect feedback from their principal and teachers for reflective purposes.

Some coaches attended monthly curriculum team meetings to learn about staff needs.

Typical Week

PLCs make up more than half the week.

Intervention takes up time although it is not on the coach pie chart.

Time is spent creating professional development materials.

Some coaches reported co-teaching and teaching in certain instances.

PLCs

Recommended: A trimester PLC cycle since there are fewer meetings at the start and end of the school year.

PLCs should focus on instruction, not just data.

Teachers often get pulled from PLCs due to parents, IEPs, etc.

It is hard to follow up with every teacher in a because there is not enough time between meetings.

In addition to this feedback and these recommendations, there was one theme that was repeated in every group: the tone set by the principal regarding instructional coaches facilitates or hinders the effectiveness of the coach in the school. Coaches also expressed the need for their own PLCs—that is, an opportunity to share, learn, and collaborate with each other. The feedback from coaches, in concert with educator feedback, formed the basis of the recommendations for PLCs below.

Recommendations

In summary, there has been growth in favorable educator perception surrounding collaboration, time spent on collaboration, use of PLCs, and teacher access to instructional resources. Teacher respondents to surveys indicated that instructional coaches have a limited impact on classroom management and knowledge of the evaluation rubric. Based on the qualitative findings in this report, the district may wish to consider the following suggestions regarding PLCs and instructional coaches:

- Expand PLC focus to include thorough coverage of the evaluation rubric and its indicators
- Include conversations about classroom management practices in PLCs
- Provide instructional coaches with training around classroom management best practices
- Ensure that intervention-related tasks do not interfere with PLC meetings
- Deliver professional development sessions separate from PLCs, so as not to interrupt or detract from PLC time
- Explore TAP practices that are viable within the TEAM structure
- Provide more collaborative time for instructional coaches to work with each other
- Consider reshaping the PLC cycle structure to accommodate beginning of year, testing, and end of year interruptions

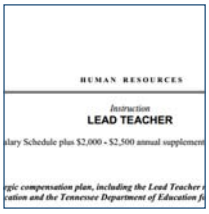
LEAD TEACHERS

Candidates

- Professional teachers with strong value-added scores, above average teacher effectiveness scores, and leadership abilities

Selection

- School-based application process
- Principals make selections from existing faculty for an annual, renewable contract.



Preparation

- Instructional rubric and observer training
- Marzano's high-yield strategies

Application

- Instructional suppport and coaching
- Classroom observations, pre- and post-conferences, critical conversations
- Create and lead professional development

The lead teacher position was created to provide additional career opportunities for teachers aside from administration. Lead teachers provide instructional support and coaching, as well as rate classroom observations in TEAM-based schools while also maintaining regular teaching duties. Lead teachers also work with school leadership and instructional coaches on staff professional development and training needs, assist with collecting and analyzing school data, lead or facilitate PLCs, and contribute to general school improvement planning. Lead teachers are meant to be instructional leaders, as opposed to operational or administrative support that is not directly related to instructional improvement and student academic growth.

All TEAM schools are allocated lead teacher positions based on the number of certified staff in the school. The typical ratio is one lead teacher for every 12 elementary teachers and one for every 15 secondary teachers. Lead teachers are selected from among existing school faculty for a yearly contract. Key qualifications for being selected as a lead teacher include demonstrated teaching effectiveness and leadership abilities. Lead teachers earn $2,000 to $2,500 in addition to their base salary annually. There were 226 lead teachers in the KCS in SY1314, with more than half in elementary schools, a third in high schools, and the remaining amount split between middle schools and alternative schools.

One of the key tasks that lead teachers assist school administrators with is classroom observations as part of the TEAM formal evaluation process. In order to conduct observations, lead teachers use substitutes to teach their classes. To track the utility of lead teachers as they impact the observation process, we considered the percent of classroom observations conducted by a lead teacher in each school. Additionally, we reviewed the number of substitute days requested to release lead teachers for observations in each school.

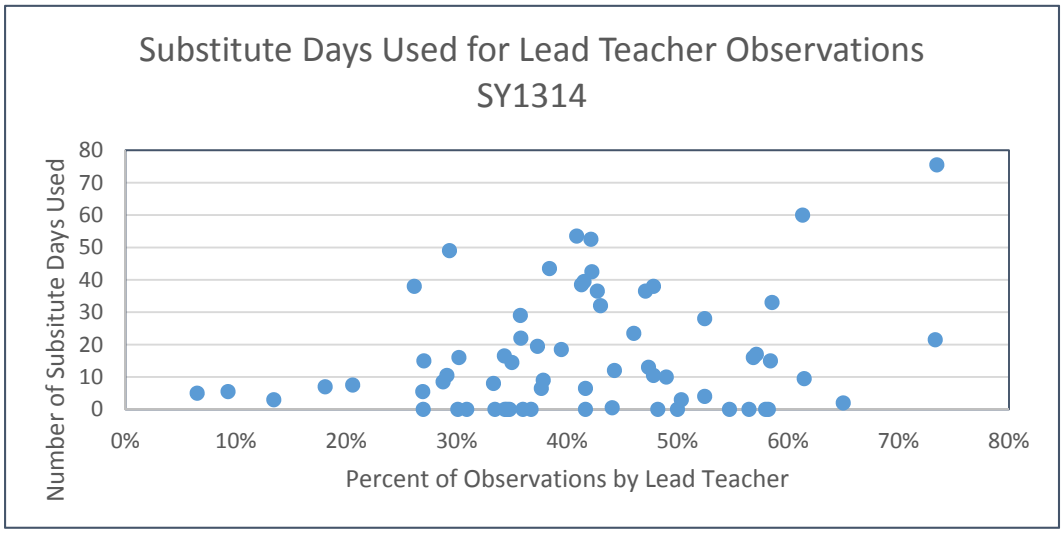
Findings

The district average percent of observations led by a lead teacher in TEAM schools was 33% in SY1314, which is down two percentage points from SY1213. The range was a minimum 6% of observations at Sequoyah Elementary to a maximum of 74% at Christenberry Elementary. The table below contains the percent of classroom observations conducted by a lead teacher in each school, ranked by largest to smallest percent, as well as the days allotted and used by lead teachers to conduct observations.

School	% of Observations by Lead Teacher	Allocated Substitute Days	Actual Substitute Days
Christenberry Elementary	74%	90	36.5
Mooreland Heights Elementary	69%	45	21.5
Farragut Primary	69%	75	75.5
Gibbs Elementary	66%	60	28
Blue Grass Elementary	60%	60	60
Bearden Elementary	60%	30	2
Beaumont Elementary	53%	30	9.5
Farragut Middle	51%	30	0
Gresham Middle	51%	36	15
Powell High	51%	48	0
Inskip Elementary	50%	45	17
Amherst Elementary	49%	60	33
Sunnyview Primary	49%	15	16
Powell Middle	48%	24	6.5
Cedar Bluff Elementary	47%	135	38.5
Halls High	46%	24	3
Farragut Intermediate	45%	90	42.5
Gibbs High	45%	18	0
Chilhowee Intermediate	45%	15	13
Shannondale Elementary	45%	30	0
Karns Middle	44%	30	4
Maynard Elementary	43%	30	10.5
Copper Ridge Elementary	41%	30	38
Rocky Hill Elementary	41%	60	52.5
Karns High	41%	42	0
Mt. Olive Elementary	41%	45	10
Powell Elementary	41%	60	36.5
Bonny Kate Elementary	40%	30	12
Bearden Middle	39%	24	23.5
Corryton Elementary	39%	30	6.5

School	% of Observations by Lead Teacher	Allocated Substitute Days	Actual Substitute Days
Sam E. Hill Family Center	39%	15	0
Fountain City Elementary	38%	30	18.5
West High	37%	42	0.5
Sterchi Elementary	37%	45	39.5
Hardin Valley Academy	37%	48	0
South Knox Elementary	35%	15	9
Halls Elementary	35%	60	53.5
Farragut High	33%	48	0
Northshore Elementary	33%	45	43.5
Fair Garden Family Center	33%	15	8
Ball Camp Elementary	31%	30	16.5
West Hills Elementary	31%	30	19.5
Halls Middle	31%	30	32
Hardin Valley Elementary	31%	90	29
Fulton High	29%	36	0
Bearden High	27%	42	0
Carter Elementary	27%	45	10.5
Central High	26%	42	0
Gap Creek Elementary	26%	15	0
A. L. Lotts Elementary	25%	45	38
Pleasant Ridge Elementary	25%	15	8.5
Adrian Burnett Elementary	24%	45	15
Karns Elementary	23%	45	49
New Hopewell Elementary	23%	30	5.5
Brickey-McCloud Elementary	23%	45	16
Richard Yoakley	18%	15	0
West Valley Middle	18%	36	22
Ridgedale Alternative	18%	15	0
Cedar Bluff Middle	18%	12	7.5
Norwood Elementary	14%	30	7
Knox Adaptive Education Center	13%	15	14.5
Whittle Springs Middle	9%	18	3
Green Magnet Elementary	7%	45	5.5
Sequoyah Elementary	6%	30	5

The figure here displays the same data in scatter plot form.



The majority of schools had roughly half of their observations conducted by a lead teacher, with varying amounts of substitute days used. There were several schools in which lead teachers conducted observations but did not use any substitute days, which raised a few questions:

- Are lead teachers able to complete observations during their plan period?
- Should lead teachers use their plan periods to complete observations?
- Are lead teachers given adequate time to perform observations?
- Are lead teachers taking adequate time to perform observations?

In a previous analysis by the Parthenon Group, survey data suggested that lead teachers are perceived by teachers to be less effective than administrators in conducting the observation process. That may be due to peer-to-peer feedback issues—teachers may not value feedback from their colleagues as highly as they would their administrators. The same study by Parthenon found that principals greatly value their lead teachers, particularly in the area of conducting observations. Does the amount or type of time lead teachers use have an impact on these mixed perceptions?

Recommendations

In addition to providing teachers with administrative experience, the lead teacher position also supports the overall instructional process at schools. The practice of using lead teachers for observations is a useful one for principals. The findings of this cursory analysis do not call for any significant changes or recommendations. Therefore, the district should continue providing appropriate training for lead teachers as it relates to the evaluation rubric. However, school administrators should review time available and time used by lead teachers for observations to ensure lead teachers are able to conduct observations without interfering with plan periods.

Strategic Compensation

To be successful, students need successful teachers and building leaders—in short, academic excellence can only be achieved with effective educators. In order to retain and compensate excellent educators, the district uses two strategic compensation plans: TAP and APEX. TAP is an evaluation system that includes a compensatory reward component. Advance, Perform, Excel (APEX) is a KCS initiative tied to the TEAM evaluation structure that rewards exceptional educators. Both TAP and APEX are used as tools to recruit, develop, promote and compensate outstanding teachers for achieving academic success.

Effective Educators

- KCS must recruit, select, induct, develop, support, promote, compensate, and retain personnel with a focus on quality and instructional excellence.
- Recognizing the impact of human capital on student achievement, KCS will be more deliberate in the development of its principal and teacher "pipelines" that will cultivate high quality leaders and strong educators for the future.

Source: Knox County Schools Strategic Plan (2009-2014)

The district invests the majority of its resources in its people. Over three-quarters of the KCS budget is dedicated to personnel costs in the form of salaries, taxes, and benefits, and teachers are the lion's share of such expenditures. (For more information, our 2012 ROI report includes a full breakdown of district spending sources and allocations.)

Investment Analysis

The investment analysis for the strategic compensation evaluation is based on the number of educators who are eligible to receive TAP or APEX reward money, not necessarily every certified staff member in the district. These costs include personnel services and benefits.

Initiative	FY14 Expenditures		Total FY14 Expenditures	# of Educators	Cost Per Teacher
	Grant	General Purpose			
TAP	$1,816,623	$1,137,238	$2,953,861	930	$3,176
APEX*	$1,773,988	$1,224,635	$2,998,623	3,689	$813
STRATEGIC COMPENSATION	**$3,590,611**	**$2,361,873**	**$5,952,484**	**4,619**	**$ 1,289**

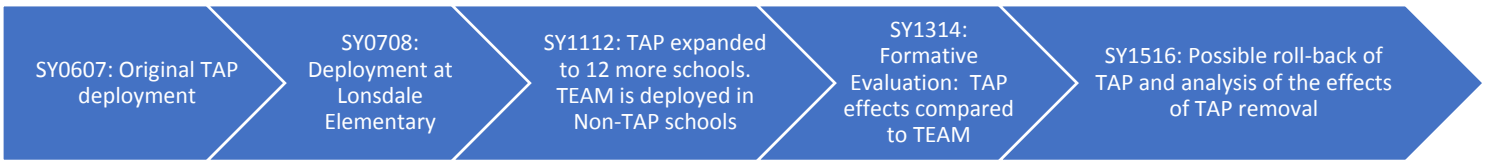
TAP

In SY0607, Knox County Schools implemented TAP: The System for Teacher and Student Advancement (previously known as the Teacher Advancement Program) in 6 of its highest needs schools. TAP is a comprehensive school reform model with the goal of increasing teacher recruitment, retention, motivation, practices, and performance. The program was expanded in SY1112 and now includes 18 schools.

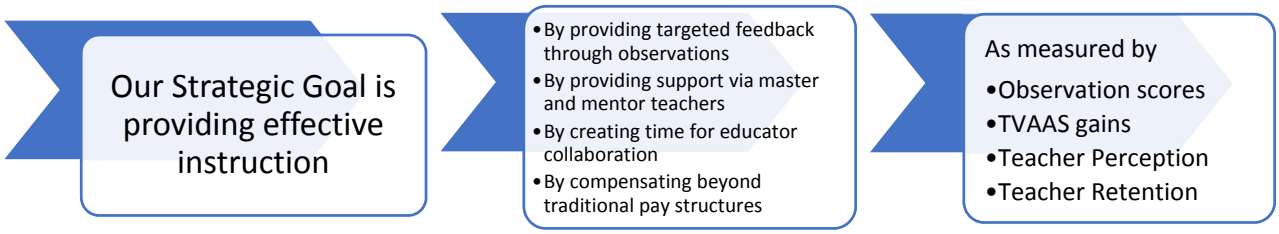
The TAP targets four specific areas as the core drivers for their theory of action.

- **Multiple career paths:** Master and mentor teachers serve as school leaders to provide input and support.
- **Ongoing professional growth:** Structures are put in place to drive teacher collaboration and interaction with master/mentor teachers in cluster meetings.
- **Instructional based accountability:** A research-based rubric serves as a vehicle for providing targeted feedback on teaching practices.
- **Performance based compensation:** Teachers are eligible for bonus pay based on observation scores and student outcomes. Supplemental pay is provided for additional leadership responsibilities.

In SY1112, non-TAP schools were exposed to some of these elements as part of the Tennessee Educator Acceleration Model (TEAM) coupled with other KCS initiatives (strategic bonuses, professional learning communities, and deployment of TEAM lead teachers). This study analyzes differences between the 12 TAP schools that joined the program in SY1112 and a pool of the most similar TEAM schools.



Findings



- **Growth on observation scores:** TAP teachers exhibited higher rates of change in observation scores in their first 4 years of teaching. This is evidence that the TAP system provides novice teachers with a better understanding of the classroom observation rubric. This finding is generally corroborated by district level practitioners that have experience with both TEAM and TAP systems.
- **Teacher perceptions:** TAP teachers are more likely to have favorable perceptions of professional development offerings and the feedback provided through the observation process. However, TAP teachers are less likely to respond that they have time to collaborate with their colleagues. It is possible that interactions between classroom teachers and master/mentor teachers are viewed as coaching sessions rather than true collaboration.
- **Teacher retention:** TEAM schools retain more teachers that meet or exceed expectations (as measured by their adjusted summative observation scores) when compared to TAP schools.

- **Student outcome data:** The analysis indicates that there is no statistical difference in mean TVAAS gains between TEAM and TAP elementary and middle schools. TAP high schools exhibit increased mean TVAAS gains when compared to their TEAM counterparts.

Recommendations

The results of the analysis are mixed. There is evidence that TAP provides a vehicle for teachers to demonstrate higher rates of growth in the quality of their instruction. However, this effect appears to be limited to early career teachers. Furthermore, those gains may not be sufficiently large to affect the student outcome data (as estimated by TVAAS) of elementary and middle schools.

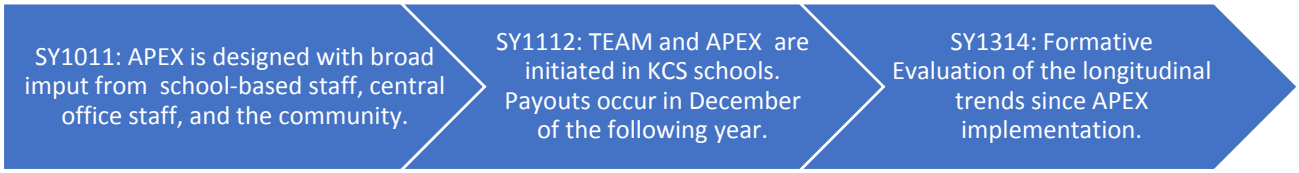
The high school data shows more promising trends. There is evidence that student outcome data (as estimated by TVAAS) is being impacted by TAP. Survey data suggests that the magnitude of change (pre-to-post TAP) was higher in high schools than elementary and middle schools. Anecdotally, high school principals are actively using TAP performance bonuses as recruitment tools. There is some evidence in the student outcome and classroom observation data that TAP schools are employing more effective first year teachers than their TEAM counterparts.

Going forward, budgetary constraints may be a barrier to the continuation of TAP. Any contraction of the TAP program will present another opportunity to estimate its true effect. Detailed analyses will occur comparing any schools that continue with the TAP program and schools were TAP is discontinued.

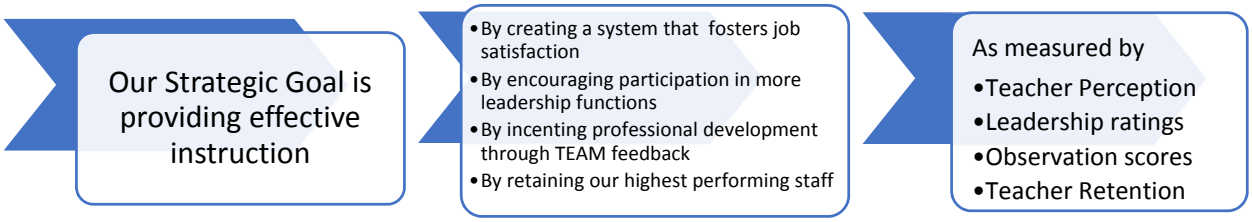
APEX

In SY1112, Knox County Schools implemented a strategic compensation system called Advance, Perform, Excel (APEX). APEX was viewed as a complementary component to the Tennessee Educator Acceleration Model (TEAM) evaluation system. Under the APEX compensation system, eligible staff members can earn up to $2,000 per year in bonus pay.

The metrics that are involved in determining the APEX bonus amount is closely tied to the APEX theory of action. APEX was designed to reward high quality instruction, student achievement, teacher leadership, and continued service in high-needs schools. As such, the APEX calculation includes TEAM observation data, a variety of student outcome metrics, leadership rubric scores, and years of continual service in the same high-needs schools.



Findings



- **Job Satisfaction:** Annual surveys suggest that the majority of KCS teachers felt that APEX has negative effects at their school. A minority of survey respondents felt that APEX has been effective at identifying effective teachers, helps teachers feel valued as professionals, and helps increase job satisfaction. Survey responses tend to shed a negative light on the APEX initiative.
- **Leadership:** There is evidence that the number of teachers that are not assuming leadership responsibilities has decreased each year APEX has been awarded (as measured on the APEX leadership rubric), and that TEAM staff are assuming more leadership responsibilities. These results are evidence that APEX is positively affecting the number of teachers that are assuming leadership responsibilities in the district.
- **Observation scores:** There is evidence that the mean observation scores of professional teachers have increased each year since APEX was initiated. Survey results are mixed regarding APEX as the driver for any gains in teaching practices. It

is possible, therefore, that gains in observation scores are more closely related to the TEAM process than they are associated with APEX.

- **Teacher Retention:** There has been no appreciable change in overall staff retention rates since APEX has been initiated. There are also no systematic trends regarding retention of teachers by effectiveness level as the APEX initiative has matured. Survey responses continually indicate that strategic compensation has little bearing on retention decisions. There is little evidence to indicate APEX has any effect on teacher retention.

- **TVAAS:** Elementary and middle schools show no real trend in TVAAS gains since the implementation of APEX. High schools exhibit a shallower trajectory that indicates yearly losses in TVAAS gains have slowed since APEX has been implemented. Survey results are mixed regarding staff perceptions on the impact of APEX on student learning.

Recommendations

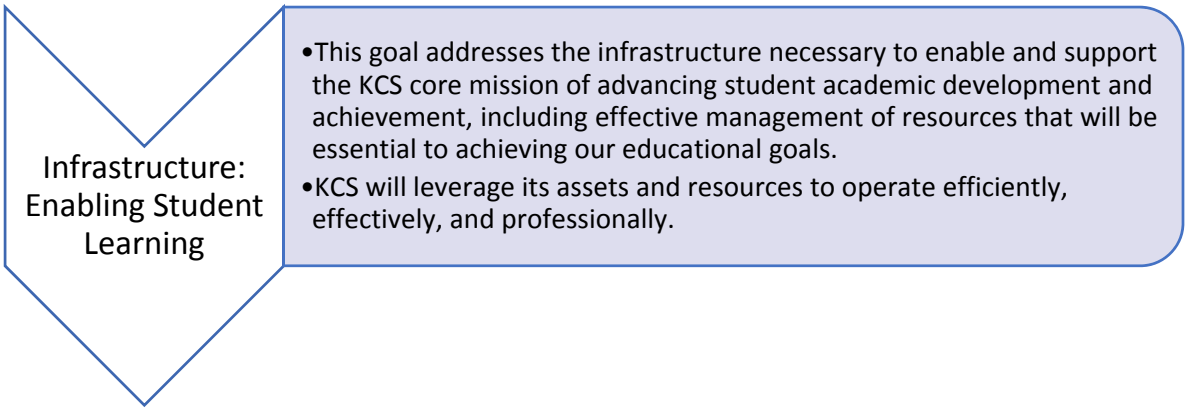
It is important to note that TEAM and APEX were launched in the same year. It will be impossible to create any causal links between outcome measures and APEX because of how intertwined it is with the TEAM process.

Survey data indicates strong negative feelings about APEX. An overhaul of the system will be required if the district is committed to continuing its strategic compensation initiative for TEAM school. Any redesigned system must determine the root cause for the negative survey responses and eradicate them from the strategic compensation model. The architects of any APEX replacement must determine if the strategic goals of APEX can even be met with a strategic compensation program. It is entirely possible that the goals of the APEX system are misaligned with the on-the-ground effects of performance-based bonuses.

Survey data does shed some light into what the majority of KCS respondents think are worthy of bonus pay. The metrics with the highest response rates for most important are service in hard to staff schools and hard to staff subject areas. KCS should consider making these two metrics key components in any revamped differentiated pay system.

Staffing Models

Recognizing that the majority of the KCS budget is designated for classroom instruction and instructional support, it is important to have a rational means of allocating personnel to individual schools. Currently, KCS uses a budget allocation formula that is transparent and rational. However, it is necessary to review and adjust the staffing model annually to ensure its purpose of academic excellence translates to rational allocations. Additionally, the scheduling structure within a school also impacts the use and placement of staff. These logistical demands of education are tied to the goal of enabling student learning.



Source: Knox County Schools Strategic Plan (2009-2014)

In an effort to consider allocation of resources, we evaluated both staffing and scheduling (and its consequent impact on staffing) at the secondary level. In particular, we sought to quantify how class size is associated with student performance. Taking it one step further, we also examined scheduling practices for SY1314 to determine any potential impact on staffing by moving from block scheduling to a traditional schedule.

Investment Analysis

Since scheduling and staffing are already in place and their costs are reflected in actual school costs, we reported investment changes in the form of potential decreases in full time equivalent (FTE) positions. It should be noted that

- the conservative model is based on the evaluation presented here by the REA and
- the idealized model is based on projections from the Finance department.

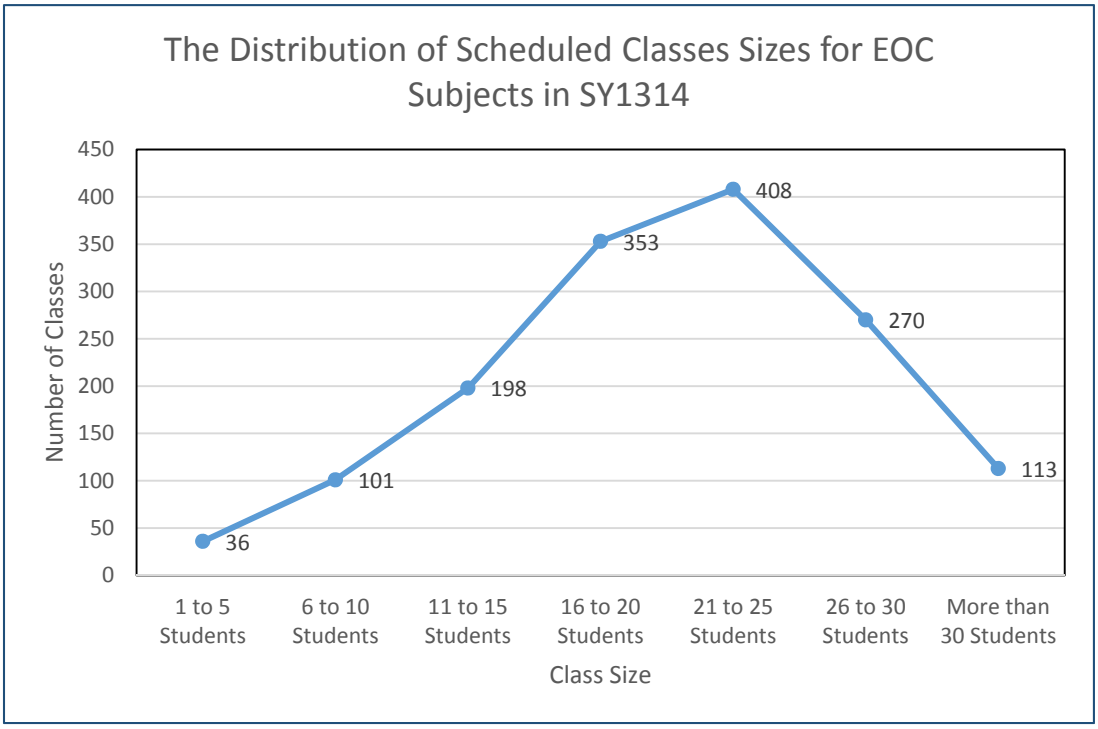
Initiative	FY14 Actual FTEs	Potential FTE Decrease	Potential FTEs with Schedule Changes
Conservative Model of Traditional Scheduling	917	41	876
Idealized Model of Traditional Scheduling	917	111	806

STAFFING RATIOS

To what extent is class size associated with student performance? This evaluation will consider the distribution of class sizes through the scheduling data for the school year 2013-2014 (SY1314). From there we will investigate student growth as measured by the difference between each student's predicted scale score and actual scale score. This data exists for the eight courses for which there are state end-of-course (EOC) exams: Algebra I, Algebra II, English I, English II, English III, Biology I, Chemistry I, and United States History. We will not be looking at student performance data due to its inherent socio-economic bias. Other demographic features such as special education status and English Language Learner status would also be grounds for an achievement bias. The growth data eliminates these biases in that the predicted scores are produced by only considering the observed scores in relationship to each student's previous test history. This means that there is a theoretical level playing field when we consider growth in this manner. Of the 30,599 high school student schedules that could have resulted in an EOC test, 27,257 students took their EOC exam and had a sufficient test history to create a predictive score and subsequent growth score.

Findings

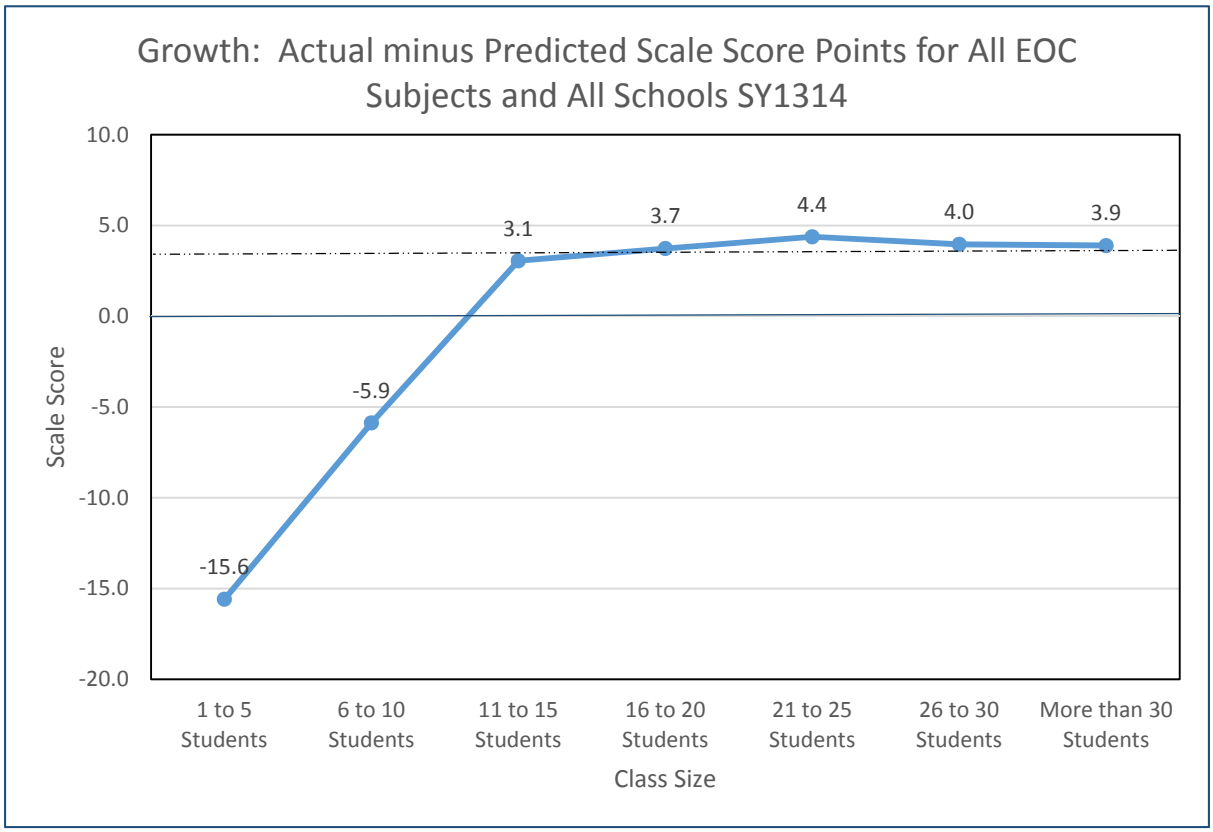
We were able to examine 1,479 scheduled classrooms. The most popular classroom size across the district for EOC courses in SY1314 was between 21 and 25 students. This can be seen in the figure below.



In terms of student growth, the mean growth of all 27,423 students over all of the schools and all of the subjects was just a little over 3.6 scale score points. These results can be found in the table and figure below.

	Mean Growth by Class Size							
	1 to 5 Students	6 to 10 Students	11 to 15 Students	16 to 20 Students	21 to 25 Students	26 to 30 Students	More than 30 Students	Total
All Subjects	-15.6	-5.9	3.1	3.7	4.4	4.0	3.9	3.6
Difference From Total	-19.2	-9.5	-0.6	0.1	0.7	0.3	0.3	0.0

The results show negative growth for the first two class size categories. This indicates that the observed mean scale score was less than the predicted mean scale score in these two categories. The opposite was the case for the remaining categories where the results vary between 3.1 and 4.4 scale score points. The green line in the figure represents the threshold for zero growth while the overall mean growth of 3.6 scale score points is represented by the dotted line. Are the means for the individual class sizes significantly different than zero? Are they significantly different from the overall mean of 3.6?



Our test results indicated that the class sizes of 1 to 5 students and 6 to 10 students performed significantly below zero, their predicted mean, as well as the overall mean growth of 3.6 scale score points. All other class sizes had means that were above zero in a statistically significant manner, while the class size of 21 to 25 students performed above the overall mean in a statistically significant way.

Recommendations

To what extent is class size associated with student performance? Overall there is a significant association. Classes with ten or fewer students had a smaller growth mean than classes of 11 or more in a statistically significant way.

It should be pointed out that association does not imply causation. While there is an association between class size and student growth, it should be pointed out that there are other considerations that go into class size. Some of these considerations that may be affecting the growth results.

- Class sizes may be intentionally smaller at some locations for non-academic reasons. A reason could include limiting size for better classroom management.
- Schedule makers may intentionally or unintentionally overload the classrooms of the school's better teachers and under fill the classrooms of teachers whom they perceive to be less effective or not as student friendly.

Outside of purposeful student scheduling, there are other possible reasons for the relationship between class size and average growth.

- In small classrooms, students may feel more vulnerable because they may not have a peer that has a similar point of view or is struggling with a concept in the same way. If this is the case, students may be less inclined to actively participate in the classroom.
- Teachers may be more invigorated by the energy that a small class of students may not be able to provide. Our data suggests that this "critical mass" may be somewhere between 21 and 25 students in the classroom.
- When a few students are absent from a small class, a teacher may not want to push ahead for the sake of the missing students whose absence will be notable. In a larger class, the teacher may be more willing to press ahead with the curriculum and expect the missing students to catch up.
- When class sizes exceed 25 students, the mean growth starts to diminish. This may be due to the additional efforts required to maintain classroom management. On a related note, the larger the class, the more difficult it is for the teacher to establish a relationship with each student and his or her parents or guardians. This limits the amount of time available for student-led performance conferences, which have been shown to positively affect student performance.

SECONDARY BLOCK SCHEDULING

In the early 1990's, Knox County Schools implemented block scheduling as the standard schedule structure in all of its high schools. For the implementation, all high schools converted from a traditional six-period school day to a 4x4 block. This change had major implications for students, teachers, and administrators. Students had to adjust to longer class periods and a faster pace because courses only met for half of the school year. Teachers had to adapt by incorporating new teaching strategies that are appropriate for a 90-minute block. In addition, teachers gained more planning and collaboration time under block scheduling. Administrators were able to offer more courses and were less constrained when creating their teacher and student schedules. This was due to the fact that teachers taught more classes over the course of the year even though they are actively teaching a smaller percentage of each day. It is for this reason that it generally takes more teachers to operate in a block scheduling scheme. The purpose of this report is to examine scheduling practices for SY1314 and to determine any potential impact on staffing by moving back to a traditional schedule.

Findings

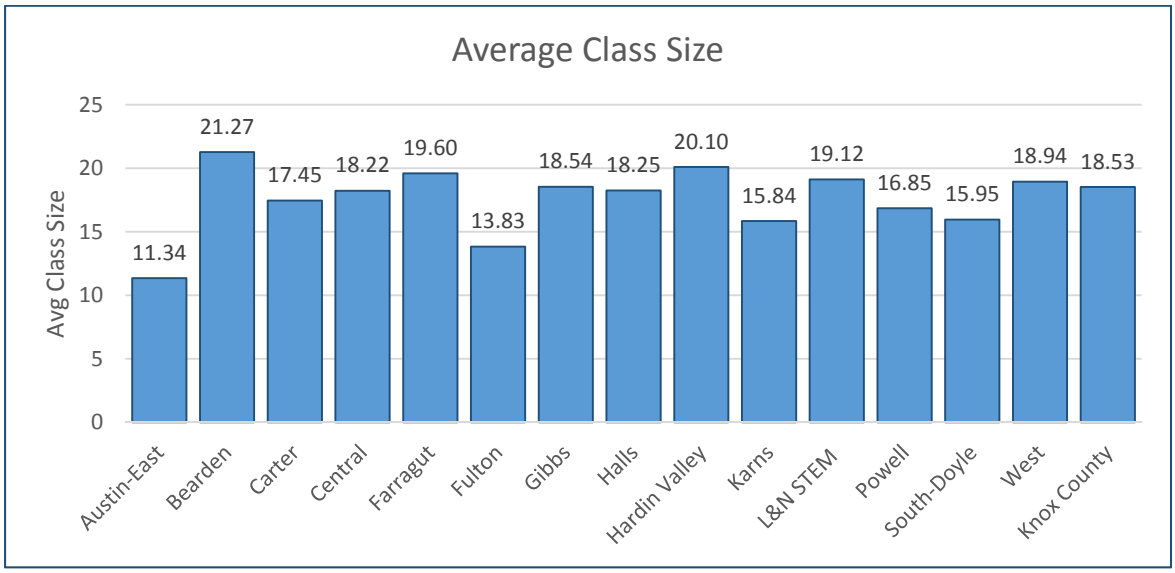
In order to determine how teaching staff were being utilized across the district, several different measures were examined: SY1314 schedule structure, class size at each school, and the proportion of possible sections that were scheduled. All but one of the active high schools in SY1314, Dr. Paul L. Kelly Volunteer Academy, were still utilizing some form of block scheduling. Since block scheduling is not in use at Kelly Volunteer Academy, it was not included in this analysis for SY1314. In all other high schools, some form of block scheduling is still in use. However, some schools have moved away from the 4x4 block scheduling structure (see the table below). Four high schools are utilizing a combination of traditional and block scheduling (modified block) for a portion of their populations and two other schools have implemented an alternating day block where classes meet every other day all year long.

School	4x4 Block	Modified Block	Alternating Day Block
Austin-East	✓		
Bearden	✓		
Carter	✓		
Central		✓	
South-Doyle		✓	
Farragut	✓		
Fulton		✓	
Gibbs	✓		
Halls		✓	
Hardin Valley	✓		
Karns	✓		
Powell	✓		
West			✓
L&N STEM			✓

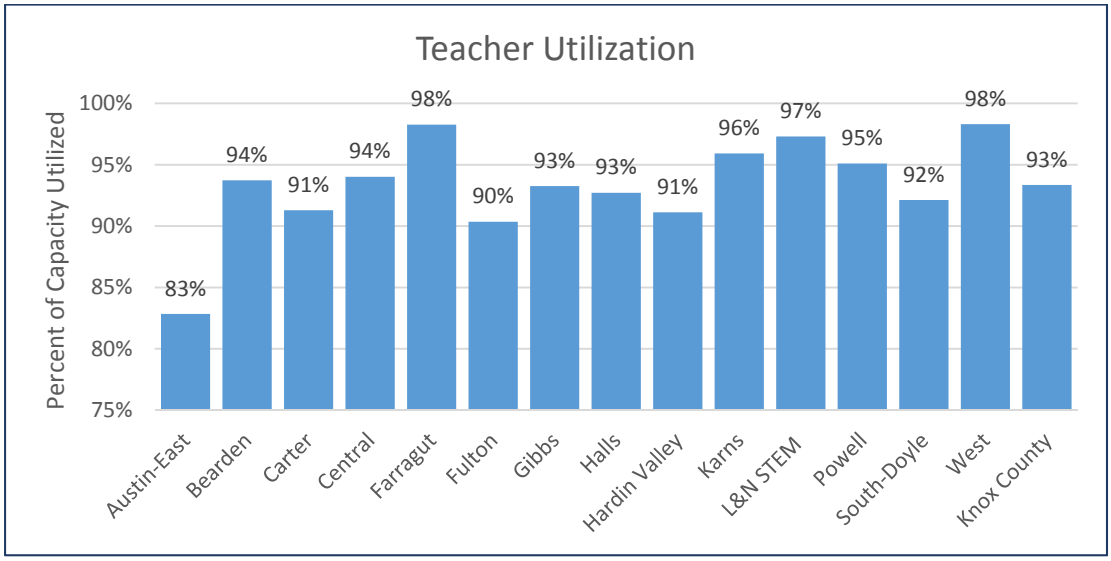
Since classes across the district are scheduled for different lengths of time, each course was converted to a standard unit of measurement. For the purposes of this analysis, the standard unit of measure is a block and it is defined as 90 minutes daily for one semester.

The notable findings from this analysis are as follows:

- The mean class size in high schools across the district was 18.53 students per class with a standard deviation of 8.97, which indicates a high degree of variation in the individual class sizes across the district. The smallest average class size by school was 11.34 at Austin East High School and the largest was 21.27 at Bearden High School (See Figure 1). In addition, 52% of all Math and ELA classes contained 20 or fewer students.



- In SY1314, the district as a whole utilized 93% of its full schedule capacity. Most schools were very close to or above the district rate with the exception of Austin East who scheduled 83% of their capacity.



- If teaching staff were utilized at the same rate and class sizes remained at the same level, then moving to a six-period traditional schedule would require approximately 41 fewer teaching positions across the district.
- Class size and teacher utilization are important considerations when projecting staffing levels. If the average class size were to increase by 5%, we project that approximately 95 fewer staff members would be required across the district when compared to block scheduling. Likewise, if it were possible to increase the average class size to 25 students per class, then 341 fewer teachers would be required.

		SY1314 Class Size		SY1314 Class Size + 5%		SY1314 Class Size = 25	
School	Teaching Staff in SY1314	Projected Staff for 6 period day	Difference	Projected Staff for 6 period day	Difference	Projected Staff for 6 period day	Difference
Austin-East	68	67	-1	64	-4	64	-37
Bearden	117	112	-5	107	-10	107	-19
Carter	68	67	-1	64	-4	64	-20
Central	78	76	-2	72	-6	72	-20
South-Doyle	89	84	-5	80	-9	80	-29
Farragut	106	99	-7	94	-12	94	-24
Fulton	83	85	2	81	-2	81	-34
Gibbs	73	72	-1	69	-4	69	-18
Halls	87	85	-2	81	-6	81	-23
Hardin Valley	122	116	-6	111	-11	111	-23
Karns	92	90	-2	86	-6	86	-32
Powell	102	97	-5	92	-10	92	-33
West	79	76	-3	72	-7	72	-19
L&N STEM	31	29	-2	28	-3	28	-8
	Total		-41		-95		-341

Recommendations

Tennessee state law provides guidelines for maximum class sizes in TCA § 49-1-104. The average number of students in non-vocational courses cannot exceed 30 students per class and no individual class can exceed 35 students. Similarly, vocational courses cannot exceed an average of 20 students per class and no individual class can be larger than 25 students. The average class size in Knox County is well below the limits imposed by state law and small class sizes across the district suggest educators are not being used to maximum effect. Since class size plays a key role in staffing, the district may wish to investigate increasing class sizes regardless of whether a change is made to the schedule structure. If a change in schedule structure from block scheduling to a traditional schedule were made, however, it would reduce the number of teaching staff required at almost every high school across the district, resulting in significant cost reductions. Additionally, a small increase in class size coupled with a traditional schedule could result in even greater cost reductions. While the potential staffing cost savings is compelling, a change in scheduling is a fundamental change in the way schools are structured and it will have far reaching implications to almost all stakeholders. Some of the possible consequences outside of staffing include:

- Students in a traditional schedule are enrolled in fewer courses when compared to a block schedule which is one of the reasons why that schedule structure requires fewer staff. Since students would not be able to take as many classes over the course of their academic career, students would have less choice and flexibility when choosing their plan of study. It is also likely that schools would be forced to limit the variety of courses that they offer.
- Since most courses in a traditional schedule span the entire year, schools would have less flexibility when building their master schedules.
- Teachers would be teaching a larger percentage of the school day and would have less time to engage in collaboration with their colleagues, planning, or participate in other school duties.
- Since all student in the district would be enrolled in their core academic subjects all year, it is likely that more textbooks would be required to operate a traditional schedule. Costs associated with buying additional textbooks would offset some of the cost savings.

Technical Reports

The following section contains the technical reports of the programs the REA evaluated. These reports offer brief descriptions of the programs, plus detailed information about the methodology used for the program evaluations. The results of our statistical analyses are presented with conclusions and considerations for any future research. The technical reports are intended for those readers who wish to understand how and why we reached the conclusions we did for each program. We also provided enough detail for any readers who want to duplicate our studies as well. Any questions about the methodology or results should be forwarded to the department at rea@knoxschools.org.

Community Schools

Prior to school year 2013-2014 (SY1314), Knox County Schools had embraced the Community School concept at four schools. Pond Gap Elementary was the flagship school, beginning Community School activities in SY1011. This initiative expanded in SY1213 to include Green Elementary, Lonsdale Elementary, and Norwood Elementary Schools. SY1314 saw the addition of Christenberry Elementary, Sarah Moore Greene Elementary, and Vine Middle Schools. This report did not examine the various services in which the various schools were engaged, but focused on the outcomes exhibited at the seven schools.

Methodology

While the entire schools were engaged with some Community School activities, the analysis focused on the roughly 725 students who actively participated in the afterschool programs throughout the year. These students were designated as Community School students and their peers as non-Community School students.

The following indicators were used in the analysis:

- Student attendance
- Discipline referrals
- Academic achievement
- Academic growth

When possible, the data was measured in two ways. As the Community School students are subsets of the schools, the first way was to measure the Community School students against their peers. The second way was to measure the Community School students against themselves where baseline data was available. For any statistical test performed, a p-value of less than .05 ($p < .05$) was considered statistically significant, as that indicated that the probability of a result that extreme happening by chance would be less than one out of twenty. The null hypothesis for any statistical test is that there is no difference between the items being tested while the alternative hypothesis is that there is a difference from which we will then consider the direction of the difference.

Results: Student Attendance

Last year saw an increase in the number of snow days and a subsequent decrease in the maximum number of days present. To achieve a degree of consistency, we prorated student absences to be out of 170. Students who were enrolled for fewer than 20 days were not considered to avoid skewing the results. While the number of students in each group is different, the distribution of absences between Community School students and non-Community School students is very similar in shape. These are presented in Figure 1.1.

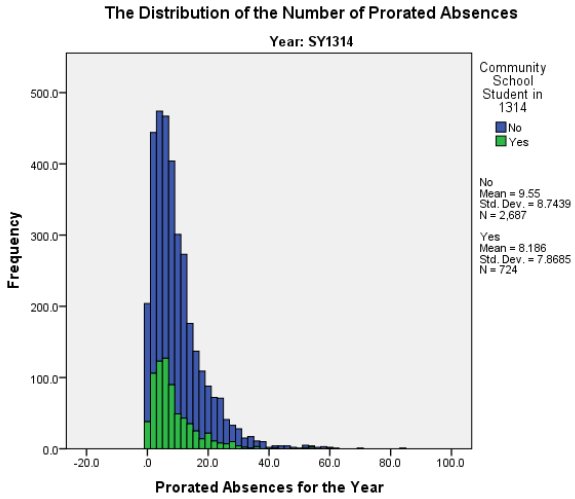


Figure 1.1: The Distribution of the Number of Prorated Absences in the Community Schools

We subjected the number of prorated absences between Community School students and their peers using a two-sample t-test for each of the schools and for the aggregate of the schools. The results of independent tests can be found in Table 1.1 below.

Table 1.1: Two-sample t-tests on the Number of Prorated Absences in the Community Schools

School	Community School Student in SY1314					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	9.4	440	7.5	122	-2.0	.010
Green	10.6	291	8.9	73	-1.7	.125
Lonsdale	7.0	326	5.9	93	-1.2	.083
Norwood	8.7	543	6.9	106	-1.8	.021
Pond Gap	9.1	241	7.7	88	-1.4	.135
Sarah Moore Greene	10.1	581	8.1	160	-2.0	.004
Vine	12.5	265	13.6	82	1.1	.409
Total	9.5	2687	8.2	724	-1.4	.000

There was a significant difference between the number of prorated absences for the two groups at three of the seven schools, as well as for the aggregate. While the number of absences was greater for the Community School students at one school (Vine Middle), over all of the seven schools the Community School students averaged 1.4 fewer prorated absences. Since students did not become Community School students through a random process, it is possible that this difference may be due to a selection bias, so causality cannot

be determined. Even so, there was a statistical association between Community School student status and the number of prorated absences. If the Community School students had the same number of absences as their peers, there would have been close to a thousand (987) more absences at these seven schools.

Baseline attendance data was gathered for 561 Community School students as well as for 1,727 non-Community School students. Baseline data was subtracted from the current year in order that a negative number would represent a decrease in the number of absences from year to year. The distribution of the change in absences is represented in Figure 1.2.

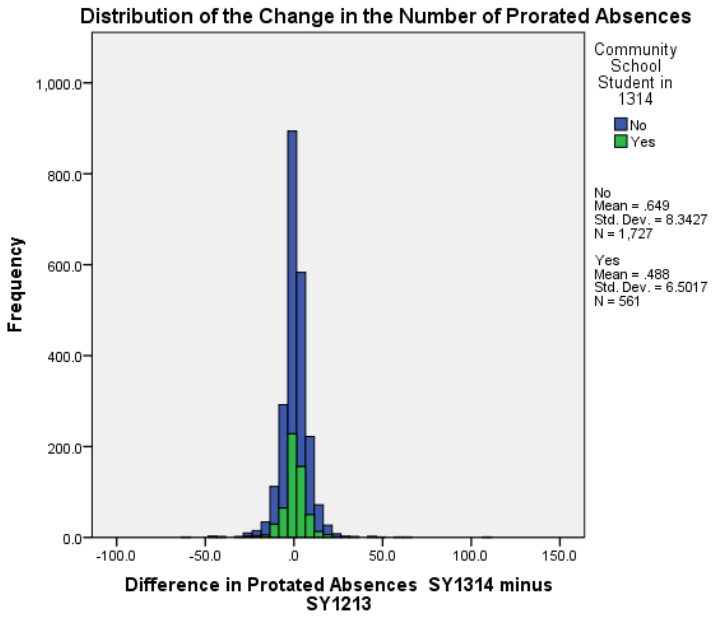


Figure 1.2: Distribution of the Change in the Number of Prorated Absences in the Community Schools

The general shape of the two groups was still the same, but they each were centered near zero. This indicated that the number of students with decreased absences was basically balanced by students with increased absences. Each group actually exhibited an increase in the number of prorated days absent in SY1314. The Community School students with two years of prorated data increased at a smaller number of prorated days than their peers by about a fifth of a day, but this difference is not statistically significant. There was one school whose Community School students saw a significant difference when compared to the non-Community School students. This occurred at Sarah Moore Greene Elementary where there was a change of close to two days difference. The data for the schools can be found in Table 1.2.

Table 1.2: Two-sample t-tests on the Change in the Number of Prorated Absences in the Community Schools

School	Community School Student in 1314					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	.2	271	1.0	101	0.8	.258
Green	-.7	195	.7	53	1.4	.223
Lonsdale	1.1	214	1.3	74	0.2	.785
Norwood	1.6	360	1.3	79	-0.3	.723
Pond Gap	1.1	145	1.7	65	0.6	.573
Sarah Moore Greene	1.9	360	.2	134	-1.7	.027
Vine	-2.5	182	-3.6	55	-1.2	.452
Total	.6	1727	.5	561	-0.2	.638

Some of the data for this subset of students was particularly interesting. Table 1.1 showed that the Community School students from Vine Middle School had a larger number of prorated absences than their peers. Yet, when just the students who had been in the Knox County Schools for each of the last two years were considered, the Community School students at Vine had an average of 1.2 fewer days absent than their peers. While this is not statistically significant, it does suggest that previous enrollment in the Knox County Schools had a beneficial effect on the Community School students at Vine Middle School.

Results: Discipline Referrals

The number of discipline referrals was prorated based upon the number of days that a student was in school in a manner that is similar to what we did with the absences. The majority of students did not have any office referrals as was the case in the 2014 EROI. But, the Community School students averaged slightly more discipline referrals than their peers. The Community School students averaged 1.4 referrals while the non-Community School students averaged one referral. This pattern was consistent across all seven schools. The difference was not statistically significant at any individual school, but the aggregate may be statistically significant. It should be noted that the distribution of prorated referrals was far from normally distributed, and as such, did not strictly qualify for being a candidate of a t-test (see Figure 1.3).

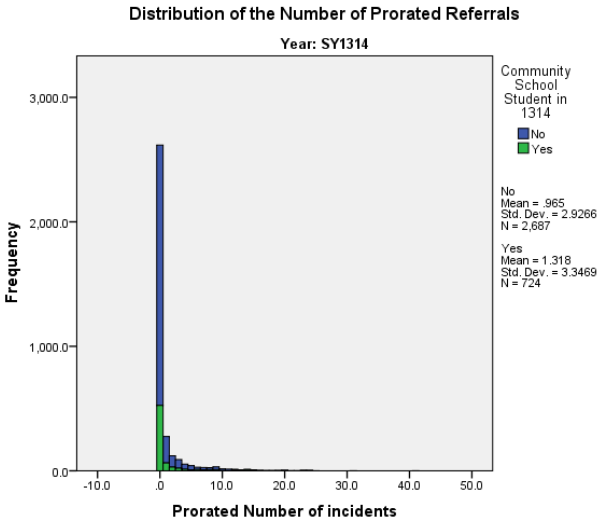


Figure 1.3: The Distribution of the Number of Prorated Office Referrals

Since the data was not normally distributed, a non-parametric test was in order. A Mann-Whitney test was applied to the data. The results are available in Table 1.3 and indicate that the total is statistically significant. This test also indicated that Norwood and Pond Gap Elementary Schools may also have had an increased number of office referrals among the Community School students. One possible reason for the larger average number of prorated office referrals would be that the Community School students' behavior is held to a higher standard. This would be difficult to verify, but remains a possibility.

Table 1.3: Two-sample t-tests and Mann-Whitney U test on the Mean Number of Prorated Office Referrals

School	Community School Student in SY1314					t-test	Mann-Whitney U test
	No		Yes		Difference		
	Mean	Count	Mean	Count	Mean	p-value	p-value
Christenberry	.2	440	.2	122	0.0	.932	.419
Green	1.0	291	1.3	73	0.3	.455	.429
Lonsdale	.1	326	.0	93	0.0	.570	.550
Norwood	.7	543	1.3	106	0.5	.060	.000
Pond Gap	.4	241	.6	88	0.3	.123	.010
Sarah Moore Greene	.9	581	1.0	160	0.1	.598	.507
Vine	4.5	265	5.9	82	1.4	.094	.194
Total	1.0	2687	1.3	724	0.4	.010	.002

We next considered the year over year difference in the prorated number of office referrals for students who were in the Knox County Schools for both years. We considered the same students that we observed for the year over year attendance. A graph of the distribution of the change in the number of prorated office referrals can be found in Figure 1.4.

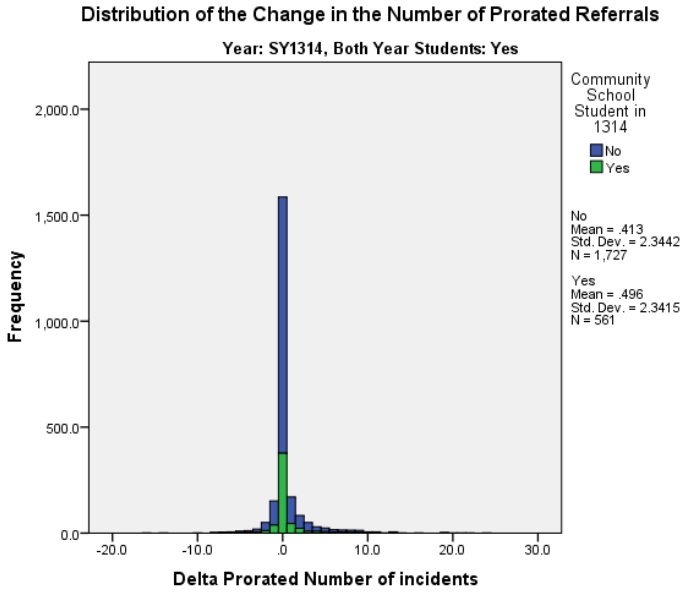


Figure 1.4: The Distribution of the Change in the Number of Office Referrals

The graph displays a very similar distribution for each type of student. Further analysis showed that the number of prorated office referrals increased by 0.4 of a referral for non-Community School students and by 0.5 of a prorated office referral for Community School students. This difference is not statistically significant in the aggregate as displayed in Table 1.4. When we considered individual schools, there was an average decrease in three schools and an average increase in the other four. None of the individual school differences were statistically significant.

Table 1.4: Two-sample t-tests on the Mean of the Change in the Number of Office Referrals

School	Community School Student in SY1314					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	.0	271	-.1	101	-0.1	0.349
Green	.6	195	.4	53	-0.2	0.658
Lonsdale	.0	214	-.2	74	-0.2	0.158
Norwood	.0	360	.5	79	0.4	0.062
Pond Gap	.0	145	.0	65	0.0	0.967
Sarah Moore Greene	.6	360	.6	134	0.0	0.970
Vine	2.1	182	3.1	55	1.0	0.145
Total	.4	1727	.5	561	0.1	0.471

Results: Academic Achievement

We examined the differences between the Community School students and non-Community School students by first looking at each group's performance on the TCAP exams in Reading/Language Arts (RLA) and Math. We were able to gather proficiency levels for 468 Community School students and 1,119 non-Community School students. For each of the subjects, the Community School students had a higher overall percentage of students who were proficient or advanced while there was variation among the individual schools. In RLA, the non-Community School students had a better overall percentage at five of the seven schools, but none in a statistically significant way. Two schools had their Community School students perform better: at Norwood Elementary the Community School students were 2.3% better while at Sarah Moore Greene Elementary they were 6.2% better and in a statistically significant way. These results can be seen in Table 1.5.

Table 1.5: Percent Proficient or Advanced in RLA along with Chi-Squared Results

School	Community School Student in SY1314				Difference	p-value
	No		Yes			
	Count	Proficient or Advanced in RLA	Count	Proficient or Advanced in RLA		
Christenberry	128	44.5%	96	39.6%	-4.9%	0.305
Green	110	14.5%	36	11.1%	-3.4%	0.630
Lonsdale	112	16.1%	35	11.4%	-4.6%	0.370
Norwood	233	29.6%	47	31.9%	2.3%	0.750
Pond Gap	98	29.6%	54	24.1%	-5.5%	0.371
Sarah Moore Greene	188	11.2%	121	17.4%	6.2%	0.047
Vine	250	21.2%	79	20.3%	-0.9%	0.784
Total	1119	23.5%	468	23.7%	0.2%	0.913

Math had a different story. In five of the schools, the Community School students performed better. There was enough evidence to conclude that overall and at Sarah Moore Greene Elementary the results were statistically significant. The achievement results can be found in Table 1.6.

Table 1.6: Percent Proficient or Advanced in Math along with Chi-Squared Results

School	Community School Student in SY1314				Difference	p-value
	No		Yes			
	Count	Proficient or Advanced in Math	Count	Proficient or Advanced in Math		
Christenberry	128	57.0%	96	65.6%	8.6%	0.099
Green	110	15.5%	36	11.1%	-4.3%	0.371
Lonsdale	112	25.0%	35	28.6%	3.6%	0.699
Norwood	233	37.8%	47	42.6%	4.8%	0.548
Pond Gap	98	30.6%	54	35.2%	4.6%	0.558
Sarah Moore Greene	188	10.1%	121	15.7%	5.6%	0.033
Vine	250	13.6%	79	11.4%	-2.2%	0.516
Total	1119	25.8%	468	30.8%	4.9%	0.015

The previous tables included all students who took the examinations. There were actually two test categories for the exams, Achievement and Modified. We were not provided with Normal Curve Equivalent scores (NCEs) for the Modified test takers, but we do have this scale variable for the majority (1,637) of students who took the Achievement tests. The previous Proficient or Advanced data looked at a categorical variable where there were only two outcomes, Yes or No. The NCE data looks at scale variables where a student can score anywhere between 1 and 99. T-tests were performed on these variables. These results can be found in Tables 1.7 and 1.8.

Table 1.7: Two-sample t-tests on the Reading/Language Arts Normal Curve Equivalents

School	Community School Student					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	49.2	123	47.8	95	-1.5	.617
Green	28.2	109	29.1	36	0.9	.809
Lonsdale	35.4	112	33.9	35	-1.6	.679
Norwood	40.8	233	43.3	46	2.4	.476
Pond Gap	39.6	98	39.7	54	0.1	.981
Sarah Moore Greene	25.4	188	33.5	121	8.1	.001
Vine	33.1	238	31.4	76	-1.7	.553
Total	35.5	1101	37.5	463	1.9	.107

In general, the results of the t-test align with the chi-squared test and again show that the Community School students performed better overall (by 1.9 NCEs), and statistically significantly better at Sarah Moore Greene Elementary. This test showed that Community School students outperformed their peers on average in four of the seven schools in RLA. This may seem like a contradiction to the Proficient or Advanced data, but it is not because the Community School students within the broad category of those who were not Proficient or Advanced, had higher scores.

Table 1.8: Two-sample t-tests on the Math Normal Curve Equivalents

School	Community School Student					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	55.3	123	59.5	95	4.2	.121
Green	31.8	109	33.6	36	1.8	.655
Lonsdale	38.5	112	41.8	35	3.3	.411
Norwood	46.8	233	49.9	46	3.1	.315
Pond Gap	43.9	98	43.8	54	-0.1	.973
Sarah Moore Greene	28.3	188	34.7	121	6.4	.004
Vine	38.9	238	33.5	76	-5.4	.046
Total	40.3	1101	42.7	463	2.4	.056

For the Math exam, the overall results did not quite reach our threshold for significance. Sarah Moore Greene again exhibited statistically significant scores that were better for the Community School students while Vine Middle School showed statistical significance in favor of its non-Community School students.

Results: Academic Growth

Each student was used as their own control in this section. Performance levels and NCEs from SY1213 were used as the baselines and growth was evaluated based on those.

We first considered the change in performance level. Students with two years of TCAP performance level data were placed into one of three change-in-performance categories: Worse, Same, or Better. It should be noted that students who had previously performed at an Advanced level could not perform at a higher level, so Same would be the desired result. In a similar manner, students who had previously performed at a Below Basic level could only achieve the Same or Better performance level. A chi-squared test was conducted on the Community School students looking at how they actually performed when compared to what they would have been expected to perform if they had performed in a manner similar to how the non-Community School students performed.

Overall, the Community School students had 3% more of their students performing better than the non-Community School students. The difference was not statistically significant. Three schools exhibited statistical significance. While Christenberry Elementary saw a slightly higher percentage of its Community School students perform at a Worse level (15.5% to 14.1% for the non-Community School students,) it also saw a very large difference in the percentage of students who performed at a Better level (34.5% to 16.9%). The other bright spot was at Pond Gap Elementary, but in a very different fashion. Community School students at the school performed Better at a slightly higher rate (11.4% to 10.6%). The striking statistics are in the percentage of students who performed Worse. The percentage of students who performed Worse was much smaller for the Community School students (11.4% to 29.8%). Lonsdale Elementary School was the one school where the change in RLA performance level was worse for the Community School students. These results for Reading/Language Arts can be found in Table 1.9.

Table 1.9: Directional Change in Proficiency in Reading/Language Arts and Chi-Squared Results

School	Change in RLA Performance Level	Community School Student					p-value
		No		Yes		Difference	
		Percent	Count	Percent	Count	Percent	
Christenberry	Worse	14.1%	10	15.5%	9	1.4%	0.001
	Same	69.0%	49	50.0%	29	-19.0%	
	Better	16.9%	12	34.5%	20	17.6%	
Green	Worse	15.1%	11	21.7%	5	6.7%	0.435
	Same	71.2%	52	69.6%	16	-1.7%	
	Better	13.7%	10	8.7%	2	-5.0%	
Lonsdale	Worse	15.1%	11	35.0%	7	19.9%	0.015
	Same	71.2%	52	65.0%	13	-6.2%	
	Better	13.7%	10	0.0%	0	-13.7%	
Norwood	Worse	12.7%	16	14.3%	4	1.6%	0.180
	Same	76.2%	96	64.3%	18	-11.9%	
	Better	11.1%	14	21.4%	6	10.3%	
Pond Gap	Worse	29.8%	14	11.4%	4	-18.4%	0.070
	Same	59.6%	28	77.1%	27	17.6%	
	Better	10.6%	5	11.4%	4	0.8%	
Sarah Moore Greene	Worse	20.7%	24	11.3%	8	-9.4%	0.094
	Same	72.4%	84	83.1%	59	10.7%	
	Better	6.9%	8	5.6%	4	-1.3%	
Vine	Worse	19.0%	44	22.9%	16	3.9%	0.566
	Same	68.1%	158	62.9%	44	-5.2%	
	Better	12.9%	30	14.3%	10	1.4%	
Total	Worse	17.6%	130	17.4%	53	-0.2%	0.286
	Same	70.3%	519	67.5%	206	-2.8%	
	Better	12.1%	89	15.1%	46	3.0%	

We next considered the change in proficiency level in Math. The chi-squared test indicated that four of the schools had statistically significant differences in the change in proficiency level when the Community School students were compared to the non-Community School students. Inspecting the changes reveals similar trends and directions for Christenberry and Lonsdale Elementary Schools. Green Elementary indicated a significant difference, but it is not in one specific direction. For this school the percentage of students who stayed in the Same level is much higher with the Community School students (78.3% to 54.8% for the non-Community School students) with smaller percentages for both Better and Worse. Sarah Moore Greene Elementary School was like Lonsdale Elementary School in that the Non-Community School students had better change in proficiency statistics than their Community School Student peers.

Table 1.10: Directional Change in Proficiency in Math and Chi-Squared Results

School	Change in Math Performance Level	Community School Student					p-value
		No		Yes		Difference	
		Percent	Count	Percent	Count	Percent	
Christenberry	Worse	11.3%	8	12.1%	7	0.8%	0.017
	Same	63.4%	45	46.6%	27	-16.8%	
	Better	25.4%	18	41.4%	24	16.0%	
Green	Worse	34.2%	25	21.7%	5	-12.5%	0.049
	Same	54.8%	40	78.3%	18	23.5%	
	Better	11.0%	8	0.0%	0	-11.0%	
Lonsdale	Worse	20.5%	15	60.0%	12	39.5%	0.000
	Same	56.2%	41	35.0%	7	-21.2%	
	Better	23.3%	17	5.0%	1	-18.3%	
Norwood	Worse	33.3%	42	32.1%	9	-1.2%	0.854
	Same	53.2%	67	50.0%	14	-3.2%	
	Better	13.5%	17	17.9%	5	4.4%	
Pond Gap	Worse	31.9%	15	25.7%	9	-6.2%	0.086
	Same	61.7%	29	60.0%	21	-1.7%	
	Better	6.4%	3	14.3%	5	7.9%	
Sarah Moore Greene	Worse	27.6%	32	41.7%	30	14.1%	0.023
	Same	63.8%	74	54.2%	39	-9.6%	
	Better	8.6%	10	4.2%	3	-4.5%	
Vine	Worse	23.6%	54	28.4%	19	4.8%	0.307
	Same	68.6%	157	68.7%	46	0.1%	
	Better	7.9%	18	3.0%	2	-4.9%	
Total	Worse	26.0%	191	30.0%	91	4.0%	0.209
	Same	61.6%	453	56.8%	172	-4.9%	
	Better	12.4%	91	13.2%	40	0.8%	

We next examined the changes in NCEs for those students who had two years of data. It is possible for a student to perform at the Same proficiency level, but have a fairly large change in NCE. For example, our RLA data showed extremes of a 28 NCE loss to a 21 NCE gain for students who had the Same proficiency level for both years. Working with the scaled NCE data, there were only two areas of statistical significance: Pond Gap Elementary in RLA, and Christenberry in Math. In each of these cases, the Community School students outgained the non-Community School students. These results can be found in Tables 1.11 and 1.12. There were other bright spots for the Community School students. Those at Pond Gap Elementary averaged a 1.2 point NCE gain in Math while the non-Community School students averaged a 1.7 point NCE loss. This 2.9 NCE difference was not enough to be considered statistically significant.

Table 1.11: Change in NCE in Reading/Language Arts and Two-Sample t-test Results

School	Community School Student					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	1.4	65	5.4	55	4.0	.108
Green	-2.8	62	-6.2	19	-3.3	.325
Lonsdale	-1.6	70	-7.5	17	-5.9	.098
Norwood	-2.2	125	0.1	27	2.4	.352
Pond Gap	-6.1	46	-0.2	33	5.9	.016
Sarah Moore Greene	-5.1	113	-3.9	71	1.2	.466
Vine	-0.4	216	-3.5	64	-3.1	.081
Total	-2.0	697	-1.6	286	0.4	.609

Table 1.12: Change in NCE in Math and Two-Sample t-test Results

School	Community School Student					p-value
	No		Yes		Difference	
	Mean	Count	Mean	Count	Mean	
Christenberry	4.8	65	9.9	55	5.1	.036
Green	-3.1	62	-0.2	19	3.0	.397
Lonsdale	1.0	70	-2.1	17	-3.1	.373
Norwood	-2.0	126	-2.0	27	-0.1	.981
Pond Gap	-1.7	47	1.2	33	2.9	.272
Sarah Moore Greene	-5.4	113	-5.8	72	-0.4	.828
Vine	-1.7	213	-4.7	63	-3.0	.096
Total	-1.6	696	-0.8	286	0.8	.397

Conclusions and Considerations

We considered the differences between the Community School students and their peers in the areas of student attendance, discipline referrals, academic achievement, and academic growth. While there were some significant differences between the groups, we could not be sure that it was not due to a potential selection bias when we considered SY1314 by itself as parents and guardians chose whether or not to opt in to the afterschool activities. We therefore also considered the changes in the various areas where the students themselves provided their own controls and then compared the differences. We considered two perspectives when looking at academic data: categorical data labels and scale score Normal Curve Equivalent data. A range of the NCEs falls into a particular category. We would prefer to use only the scale score data, but we recognize that much of the data is reported by category.

For the most part, we conducted various chi-squared tests on the categorical data and t-tests on the scaled score data. We considered a test to be statistically significant when the p-value was less than .05 which indicated that the probability of a result this extreme happening by chance was less than 1 in 20. Table 1.13 contains the areas where we found statistical significance.

Table 1.13: Areas of Statistical Significance by Student Category

Metric	CS Students Performed Better	Non-CS Students Performed Better
SY1314 Absences	Christenberry Norwood Sarah Moore Greene Overall	
Change in Absences	Sarah Moore Greene	
SY1314 Discipline Referrals		Norwood Pond Gap Overall
Change in Discipline Referrals		
SY1314 RLA Academic Achievement (Proficient or Advanced)	Sarah Moore Greene	
SY1314 Math Academic Achievement (Proficient or Advanced)	Sarah Moore Greene Overall	
SY1314 RLA Academic Achievement (NCEs)	Sarah Moore Greene	
SY1314 Math Academic Achievement (NCEs)	Sarah Moore Greene	Vine
RLA Academic Growth (Proficiency Level)	Christenberry Pond Gap	Lonsdale
Math Academic Growth (Proficiency Level)	Christenberry	Lonsdale Sarah Moore Greene
RLA Academic Growth (NCEs)	Pond Gap	
Math Academic Growth (NCEs)	Christenberry	

For SY1314, the Community School students at Sarah Moore Greene Elementary stand out as being exemplary, but this is somewhat confounded when we consider the changes from the previous year and may be subject to a selection bias. When looking at the changes from the previous year where a selection bias cannot have an effect, the Community School students at Christenberry Elementary and Pond Gap Elementary stand out, while their peers look better at Lonsdale Elementary School. In schools where the proportion of Community School students is smaller, more evidence is required to reach the threshold of statistical significance. This is part of the reason why neither group is showing a significant difference at some of our schools.

One last way to compare the two groups of students would be to apply a grading scale to their average academic growth. Using a scale similar to what the state uses for its report card grades, we get the results seen in Table 1.14.

Table 1.14: Grades Applied to Changes in NCE

School	Community School Student			
	No		Yes	
	RLA	Math	RLA	Math
Christenberry	B	A	A	A
Green	F	F	F	C
Lonsdale	D	B	F	F
Norwood	F	D	C	F
Pond Gap	F	D	C	B
Sarah Moore Greene	F	F	F	F
Vine	C	D	F	F
Total	F	D	D	D

Using this representation, the Community School students had better grades in six cells, the same grades in five cells, and worse grades in five cells. Four of the five worse grades came from Lonsdale Elementary School and Vine Middle School. Pond Gap was the one school where the Community School students had better grades in both Reading/Language Arts and Math. Congratulations are also in order at Christenberry Elementary School whose Community School students received an A in each of the subjects.

Early Literacy Support

In our previous EROI report, the SY1213 literacy intervention results were disappointing overall. One group that showed some positive intervention results came from the population for whom the intervention was originally intended, that was students in grades 1-5 who scored between the 11th and the 24th percentiles on the AIMSweb CBM reading (R-CBM) data. For SY1314, students at all 49 elementary schools were chosen for the intervention based on multiple data points that included the Spring 2013 TCAP percentile and the Fall 2014 administration of the R-CBM. Students in grades 1-5 who scored between the 11th and the 24th percentiles were still intended to be the intervention participants in conjunction with the forthcoming response to instruction and intervention (RTI^2) initiative. The intervention consisted of students receiving an additional 30 minutes of reading instruction with Voyager Passport. Classroom teachers and instructional assistants were to provide the instruction. Voyager Passport is an intervention program intended to assist students with word study, fluency, comprehension, vocabulary, writing, listening and speaking.

Methodology

We examined who participated in the intervention when using screening criteria from both R-CBM and TCAP. We focused primarily on word study, comprehension, and vocabulary as measured by the SAT-10 and TCAP assessments. We examined the effects of the Voyager intervention using the SY1314 SAT-10 predicted and observed scaled scores for grades 1-3 (using the Spring 2014 SAT-10 for grades 1-2 and the Spring 2014 TCAP for grade 3). For grades 4-5, we used the SY1213 Normal Curve Equivalent (NCE) Reading/Language Arts (RLA) score as the predicted score and the SY1314 RLA NCE as the observed score. The evaluation data also included R-CBM data and intervention participation status.

Our intent was to test Voyager student growth as measured by the difference between the observed scores and the predicted scores. Our previous report did this using three separate measures: SAT-10 scaled scores for grades 1-2, TCAP Achievement scaled scores in grade 3, and TCAP NCEs in grades 4-5. For this report, we did not combine any grades and considered each grade separately.

We considered multiple lines of inquiry in our Voyager evaluation. We tested the intervention using independent samples t-tests. For our hypothesis testing, we considered a result to be significant if the probability of a result of this kind happening by chance was less than 1 in 20 (or $p < .05$). Matched-pair analyses was also used to compare students with similar demographic and prediction features where only one of whom participated in the intervention. In an attempt to create a legitimate comparison between Voyager and non-Voyager students, we paired students based upon their demographic information and their predicted TCAP/SAT-10 Reading scores. The demographic information we used consisted of each student's gender, ethnicity, economic status, special education status, and their English language learner status. Their predicted reading scores did not have to be identical,

but did have to be fairly close –within about five scaled score points or three NCEs. These predicted scores were obtained by taking the range of observed scores and dividing it by 30.

Results: Participation

Over 20,000 students in grades 1-5 had exams scores as well as prediction scores. Over 93% of these had a TCAP RLA percentile and an R-CBM percentile. Each of these students were placed in both a TCAP and R-CBM category. These categories are listed in Table 2.1.

Table 2.1: Percentile Category Labels

Category	Percentile Range
Unknown	NA
Below Target	1 to 9
On Target	10 to 24
Above Target	25 to 49
Well Above Target	50 to 74
Extremely Above Target	75 to 99

Using these categories, the number of students in each category as well as their Voyager participation status can be found in Table 2.2.

Table 2.2: Intervention Participation with Target Classification Counts

Students In Voyager		Count of Intervention Target Using RLA TCAP/SAT-10					
		Unknown	Below Target	On Target	Above Target	Well Above Target	Extremely Above Target
		Percentile Unknown	1st – 9th Percentile	10th to 24th Percentile	25th to 49th Percentile	50th to 74th Percentile	75th to 99th Percentile
Intervention Target Using R-CBM	Unknown	16	30	26	26	9	2
	Below Target	14	216	164	75	12	1
	On Target	29	141	466	448	148	20
	Above Target	12	35	153	354	149	24
	Well Above Target	6	3	24	80	62	19
	Extremely Above Target	4	0	6	12	17	32
Students Not in Voyager		Count of Intervention Target Using RLA TCAP/SAT-10					
		Unknown	Below Target	On Target	Above Target	Well Above Target	Extremely Above Target
		Percentile Unknown	1st – 9th Percentile	10th to 24th Percentile	25th to 49th Percentile	50th to 74th Percentile	75th to 99th Percentile
Intervention Target Using R-CBM	Unknown	242	92	110	184	152	130
	Below Target	29	724	344	147	26	7
	On Target	46	248	547	547	178	29
	Above Target	62	113	491	1506	1207	441
	Well Above Target	73	26	131	974	1982	1602
	Extremely Above Target	69	1	16	246	1101	3914

Schools were given some freedom when selecting students for the Voyager intervention. Table 2.2 indicates that other considerations may have been applied. For those students who were in the 10th to 24th percentiles on both TCAP RLA and R-CBM (*On Target*), 466 were in the intervention while 547 were not. Thirty-two students who were *Extremely Above Target* on both the TCAP RLA and R-CBM participated in the intervention.

When just those within the intervention were considered, we scaled down the categories into four basic groups:

- Both Unknown or Below Target (blue)
- At Least One on Target (green)
- Both Above Target (pink)
- Conflicting Target Information (yellow)

These can be found in Table 2.3, where we noted that over 57% of those students that participated in the intervention had at least one of the metrics indicating that they were *On Target* to receive the intervention.

Table 2.3: Intervention Participation with Generalized Categories

Literary Intervention Target CBM	Intervention Target TCAP					
	Unknown	Below Target	On Target	Above Target	Well Above Target	Extremely Above Target
	Table N %	Table N %	Table N %	Table N %	Table N %	Table N %
Unknown	.6%	1.1%	.9%	.9%	.3%	.1%
Below Target	.5%	7.6%	5.8%	2.6%	.4%	.0%
On Target	1.0%	5.0%	16.4%	15.8%	5.2%	.7%
Above Target	.4%	1.2%	5.4%	12.5%	5.3%	.8%
Well Above Target	.2%	.1%	.8%	2.8%	2.2%	.7%
Extremely Above Target	.1%	0.0%	.2%	.4%	.6%	1.1%
Both Unknown or Below Target: 9.7%						
At Least one On Target: 57.3%						
Both Above Target: 26.4%						
Conflicting Target Information: 6.5%						

Results: Initial t-test

We broke the students into the generalized categories found in Table 2.3 and applied an independent samples t-test on the growth of the students comparing those who participated in the intervention and those who did not. We focused on the group that had at least one metric indicating that the student was *On Target* for the Voyager intervention. These results are in Table 2.4 while the results of all of the groups can be found in Appendix A.1.

Table 2.4: t-test Results on Growth for Students with At Least One Screener Indicating They Were *On Target* for the Intervention

			At Least One On Target			
			Predicted TCAP/SAT 10 Score	Observed TCAP/SAT 10 Score	Growth	
			Mean	Mean	Difference of Means	Count
Grade 1	Literacy Intervention Student	No	537.1	529.1	-8.0	592
		Yes	538.5	528.1	-10.4	260
	Difference (Yes Minus No)		1.4	-1.0	-2.4 (p-value=0.290)	
Grade 2	Literacy Intervention Student	No	573.6	569.7	-4.0	468
		Yes	576.3	573.2	-3.1	308
	Difference (Yes Minus No)		2.7	3.6	0.9 (p-value=0.611)	
Grade 3	Literacy Intervention Student	No	720.9	717.7	-3.2	484
		Yes	728.4	719.5	-8.9	396
	Difference (Yes Minus No)		7.5	1.9	-5.7 (p-value=0.000*)	
Grade 4	Literacy Intervention Student	No	35.2	32.3	-2.9	578
		Yes	36.9	34.1	-2.8	350
	Difference (Yes Minus No)		1.7	1.8	0.1 (p-value=0.874)	
Grade 5	Literacy Intervention Student	No	33.6	31.6	-2.0	565
		Yes	34.2	31.3	-2.9	311
	Difference (Yes Minus No)		0.6	-0.3	-0.9 (p-value=0.349)	

*notates statistical significance

Table 2.4 shows mixed results within the Reading portions of the TCAP and SAT-10 for students who were in intervention when compared to their peers who were not. In grades 2 and 4, the intervention students exhibited a (directionally) higher mean growth, but the difference was not statistically significant (α=0.05). In grades 1 and 5 the intervention students exhibited a (directionally) lower mean growth, but again the difference was not statistically significant. In grade 3, the difference in mean growth for students in the intervention was statistically different students who were not in the intervention. Visual inspection indicated that students who were in Voyager exhibited lower mean growth on the TCAP than students who were not in Voyager. The confounding aspect of this table lay in the

predicted scores for the students who were chosen for intervention. In every grade, the students chosen for intervention had a higher mean predicted TCAP/SAT-10 RLA score than the students who did not participate in the intervention. It was surprised that the students considered as On Target for intervention had higher mean prediction scores than the non-intervention students. This provides further evidence that a variety of methods were used to determine which students received the Voyager intervention.

Qualitative Follow-Up

After considering the overall results, we computed the growth for all students by school and grade and according to each student's intervention status. These results are available in Appendix A.2. We noted a few schools where the mean growth of the intervention students indicated they were performing better than their peers. In an effort to understand the practices at these schools, we met with some of the people involved in the schools' intervention programs. The overall impression from these schools was that they did not rely on Voyager as a scripted entity, but adapted it to fit their needs while applying common best practices. The following includes some of the practices that these schools implemented:

- Set apart intervention time as instructional time only so as not to be interrupted
- Used a school-wide intervention time so no stigma was attached to a student
- Started the day with intervention while students were fresh
- Made sure the intervention staff had a relationship with the students
- Chose trusted staff to be the interventionists
- Used the intervention for the intended percentiles (11^{th} to the 24^{th})
- Made team decisions on students exiting the intervention that included the teacher and did not rely entirely on Voyager indications
- Kept the intervention groups small
- Followed the program, but did not necessarily read the script
- Kept it going at a fast pace and used additional techniques that were not in the Voyager curriculum
- Charted student progress so students could monitor their work
- Kept students excited about learning

Results: Matched-Pair Analysis

Our matched-pair methodology allowed us to match 4,988 students among the five grades. If all of the students who qualified for the intervention through either TCAP/SAT-10 RLA or R-CBM percentiles were in the intervention we would not have had a suitable number of comparison students. As this was not the case, we had a sufficient number of students to create a control group of students who qualified for intervention but were not enrolled in the program. The results of the independent samples t-tests using our matched-pair design are in Table 2.5.

Table 2.5: t-tests results on the matched-pairs based on demographics and predicted scores

			Predicted TCAP/SAT-10 RLA Score	Observed TCAP/SAT-10 RLA Score	Growth and p-value	
			Mean	Mean	Difference of Means	Count
Grade One	Literacy Intervention Student	No	543.1	540.7	-2.5	444
		Yes	543.1	531.4	-11.7	444
	Difference (Yes Minus No)		0.0	-9.3	-9.3 (p-value=0.000)*	
Grade Two	Literacy Intervention Student	No	583.1	581.9	-1.2	484
		Yes	583.1	578.8	-4.3	484
	Difference (Yes Minus No)		0.0	-3.1	-3.1 (p-value=0.037)*	
Grade Three	Literacy Intervention Student	No	733.4	732.8	-.5	607
		Yes	733.4	726.2	-7.1	607
	Difference (Yes Minus No)		0.0	-6.6	-6.6 (p-value=0.000)*	
Grade Four	Literacy Intervention Student	No	40.7	39.7	-1.1	513
		Yes	40.7	38.2	-2.6	513
	Difference (Yes Minus No)		0.0	-1.5	-1.5 (p-value=0.055)	
Grade Five	Literacy Intervention Student	No	39.1	37.8	-1.3	446
		Yes	38.9	35.5	-3.5	446
	Difference (Yes Minus No)		-0.1	-2.3	-2.2 (p-value=0.014)*	

*notates statistical significance

The predicted mean TCAP/SAT-10 RLA scores for the two groups were almost identical. This suggests that the difference between the two groups is whether or not they participated in the Voyager intervention (along with some unobservable, non-cognitive characteristics). The results were not encouraging for the Voyager intervention students. In four of the five grades, the mean growth in the TCAP/SAT-10 RLA scores of the intervention students was significantly below that of their peers. In the remaining grade (4^{th}), the result was on the verge of statistical significance using our threshold of $p < .05$. It is difficult to fathom an intervention where participation results in poorer performance. Either this is the case, or there are other variables that have not been captured. It has already been mentioned that the selection methods for intervention participation is a potential confounding variable. It

is possible that the overarching goal of the Voyager program is not best measured by TCAP/SAT-10 scores. Voyager is an intervention that is aimed at helping students decode words and increase reading fluency, whereas TCAP and SAT-10 is testing reading comprehension and other reading standards. It is also possible that student behavior may be reflected in the results and will be an item for further consideration in any future studies.

Conclusions and Considerations

KCS uses Voyager Passport as a reading intervention in an effort to improve early literacy and increase student performance on the reading portion of state examinations. Mean growth, which we measured as the difference between the observed and the predicted score, was negative for all grades in SY1314 for KCS students as well, as for students across the state. This may be because the nature of the Reading/Language Arts tests was adjusted for Spring 2014, with a change in the reading passages as well as the number of questions. This change did not affect our ability to compare student growth based upon intervention status.

Prior to testing how the intervention students performed, we investigated a student's intervention status with respect to two metrics: TCAP/SAT-10 RLA percentile and R-CBM percentile. While the program was intended to be used with students in the 10th to the 24th percentile, we found that more students from this percentile band were not in the intervention than students who did participate in Voyager. Confounding the results further was the fact that for every grade level, the students who were assigned to the intervention had a higher predicted score than those who were not. When Voyager and non-Voyager students were tested against one another as a whole, the growth was mixed. In only one grade was there a statistically significant difference in the means. That occurred in third grade were the non-Voyager students had a growth scaled-score mean that was 5.7 points better than that of their Voyager peers.

We examined growth by school and visited some of the schools where the intervention groups performed better than their peers. In general, these schools did not feel obliged to follow the intervention guidelines strictly, but chose to incorporate them with other best practices.

In an effort to remove as much potential bias as possible, a matched-pair test was conducted between predicted-score and demographic-equivalent students. With a very large sample of equivalent students, the non-Voyager students outgained the Voyager students significantly in four of five grades. These results are very similar to what we observed the previous year. Either the intervention or how it is implemented is having an adverse effect on students' Reading/Language Arts scores, or other confounding variables are affecting the scores. We have noted that the selection process for placing students in intervention is not consistent across the district, as various subjective methods are being used along with the screeners. It is possible that this is part of the problem, but the matched-pair results indicate that the Voyager Passport reading intervention is not having a positive effect on the district's Reading/Language Arts scores.

Additional Elementary Reading Support

Twenty schools were provided with an instructional assistant to improve Reading/Language Arts scores and to help facilitate a reading intervention with designated students. While the program was initially designed for grades three to five, we found over 100 students in grades one and two who also received assistant lead intervention. This analysis is a smaller version of the Early Literacy Materials Report with a focus on the students supported by the Additional Elementary Reading Support (AERS) interventionists.

Methodology

The interventionists maintained logs of the students assigned to them. Some schools incorporated dynamic scheduling and altered the person providing the intervention during the year between a certified teacher and the AERS interventionist. From the intervention logs we found 755 students who were at one point assigned to the AERS interventionist. We linked state assessment scores to the predicted scores for the 604 students who took the Spring 2014 state assessment and had a sufficient test history to generate a predicted score. As the state assessments differ slightly from grade to grade, we examined each grade separately.

We defined growth to be the difference between the observed score and the predicted score and then used t-tests to compare the growth means between the AERS students and other students. We first compared the students as a whole for the 20 schools. We then generated control groups based upon demographic features and predicted scores and evaluated the AERS students using matched pair designs.

Results: Initial t-test

We first examined the predicted, observed and growth means by grade for all of the students at the AERS schools. These results are in Table 3.1.

Table 3.1: Predicted, Observed, and Growth Scores at the AERS Schools

	AERS Student							
	No				Yes			
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth	
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count
Grade One	568.8	565.8	-3.0	1611	532.5	511.7	-20.7	58
Grade Two	601.9	601.5	-.3	1566	569.0	562.1	-6.9	43
Grade Three	745.7	744.8	-.9	1455	729.0	722.2	-6.9	220
Grade Four	52.6	51.3	-1.3	1531	35.5	32.0	-3.5	157
Grade Five	53.4	48.3	-5.1	1464	34.4	31.3	-3.1	126

A brief inspection noted that the predicted and observed scores for students with the AERS interventionists had smaller means than the other students in the school. Some of these other students may have been in an intervention with another teacher, but we do not have any intervention logs for the certified staff. It is probable that the majority of these other students did not need any literary intervention. We then generated the differences between the two groups and ran a t-test on the results. These results are in Table 3.2. The predicted and observed difference means generated p-values that rounded to 0.000 for each grade. Since they were all the same, they were not noted in the table.

Table 3.2: t-tests on the Predicted, Observed, and Growth Differences at the AERS Schools

	Differences			
	Predicted Score Mean	Observed Score Mean	Growth Mean	p-value
Grade One	-36.3	-54.1	-17.7	0.000*
Grade Two	-32.9	-39.4	-6.5	0.070
Grade Three	-16.7	-22.6	-5.9	0.000*
Grade Four	-17.1	-19.3	-2.2	0.041*
Grade Five	-19.0	-17.1	1.9	0.120

*notates statistical significance

The growth difference means had a greater range of p-values. For this study, statistical significance was attributed when the p-value was less than .05 which indicated that the probability of a result this extreme happening by chance was less than 1 in 20. While the predicted difference mean was significantly smaller at each grade level for the AERS students, the growth difference mean was only significantly different in three of the grades. In grade five the AERS intervention students had a greater growth mean than their peers, although not in a statistically significant manner. In the other grades the AERS students had smaller growth means and significantly so in three of those four.

We broke down the difference growth means by school to look for any patterns. These results can be found in Table 3.3. Appendix B contains numerical counts and other data for these schools.

Table 3.3: Growth Difference Means by Schools and Grades

	Mean Growth Differences by Grade				
	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Adrian Burnett	-40.6	-16.0	-8.3	-5.0	3.2
Amherst	-23.7		-15.8	5.6	6.3
Ball Camp	-14.5	12.7	-2.5	5.0	-1.6
Bonny Kate			-13.9	2.8	-0.3
Chilhowee			-13.0	1.7	-2.2
Christenberry			-6.4	-14.6	8.9
Copper Ridge	-10.2		-6.8	-11.2	-6.1
Dogwood	-41.1	13.3		-2.7	-2.2
Fountain City			-5.8		
Gibbs	-4.3		4.2	4.3	7.2
Green		-7.8	4.7	3.5	0.2
Halls	-32.0	-10.9	-5.4	0.2	9.8
Inskip	-17.2		0.1		-7.5
Karns			-15.6	-12.4	-3.6
Lonsdale	-20.2		4.1	-4.7	
Norwood	-28.4		-6.6	0.9	4.4
Pond Gap			-3.3	-0.9	-0.1
Sarah Moore Greene	-13.5	-5.8	-6.8	-1.6	0.4
Spring Hill			-17.2	-1.0	4.8
West Haven	-0.2	2.5	-0.4	-1.8	5.7
Total	-17.7	-6.5	-5.9	-2.2	1.9

In two of our schools, Gibbs Elementary and Green Elementary, the AERS student growth means were greater than that of their peers in three grades while another four schools exhibited this trait in two grades.

While these results are interesting, it is probably not fair to compare these intervention students to all of the other students in their grades at the twenty schools. We therefore compared them to students with similar predicted scores through various matched-pair analyses.

Results: Matched-Pair Analysis

In an attempt to create a tight comparison between AERS and non-AERS students, the students were paired using three types of control groups. Each of them have similar predicted scores. They are as follows:

- Using students from the same schools without regards to demographics (n=505)

- Using students from all schools with the same gender, socio-economic status, special education status, and English language learner status (n=560)
- Using students from the same school with the same gender, socio-economic status, special education status, and English language learner status (n=209)

Each of these groupings has its own advantages and disadvantages. The more features that are added to the similar predicted scores mean that fewer control candidates are available. The t-test results for each of these are in Tables 3.4, 3.5, and 3.6.

Table 3.4: t-tests on Growth for Student with Similar Predictions within Each School

			Predicted Score	Observed Score	Growth and p-value	
			Mean	Mean	Difference of Means	Count
Grade One	AERS Student	No	534.5	538.9	4.4	51
		Yes	534.2	514.1	-20.1	51
	Difference (Yes Minus No)		-0.3	-24.9	-24.6 (p-value=0.000)*	
Grade Two	AERS Student	No	569.0	566.3	-2.7	43
		Yes	569.0	562.1	-6.9	43
	Difference (Yes Minus No)		0.0	-4.2	-4.2	
Grade Three	AERS Student	No	732.7	732.4	-.2	174
		Yes	732.5	724.6	-7.9	174
	Difference (Yes Minus No)		-0.1	-7.8	-7.7 (p-value=0.001)*	
Grade Four	AERS Student	No	36.7	36.3	-.4	141
		Yes	36.6	32.8	-3.8	141
	Difference (Yes Minus No)		-0.2	-3.5	-3.3 (p-value=0.027)*	
Grade Five	AERS Student	No	36.5	33.1	-3.5	96
		Yes	36.5	32.5	-4.0	96
	Difference (Yes Minus No)		0.0	-0.6	-0.6	

*notates statistical significance

Table 3.5: t-tests on Growth for Student with Similar Demographics and Predictions among All Schools

			Predicted Score	Observed Score	Growth and p-value	
			Mean	Mean	Difference of Means	Count
Grade One	AERS Student	No	533.6	521.4	-12.1	54
		Yes	533.3	513.5	-19.7	54
	Difference (Yes Minus No)		-0.3	-7.9	-7.6	
Grade Two	AERS Student	No	570.4	567.6	-2.8	43
		Yes	569.7	561.3	-8.4	41
	Difference (Yes Minus No)		-0.7	-6.3	-5.6	
Grade Three	AERS Student	No	729.8	728.2	-1.6	206
		Yes	729.5	721.7	-7.8	206
	Difference (Yes Minus No)		-0.3	-6.5	-6.2 (p-value=0.006)*	
Grade Four	AERS Student	No	35.6	36.1	.4	150
		Yes	36.3	32.4	-3.9	144
	Difference (Yes Minus No)		0.7	-3.6	-4.3 (p-value=0.003)*	
Grade Five	AERS Student	No	34.8	36.8	2.0	115
		Yes	34.7	31.7	-2.9	115
	Difference (Yes Minus No)		-0.1	-5.1	-5.0 (p-value=0.006)*	

Table 3.6: t-tests on Growth for Student with Similar Demographics and Predictions within Each School

			Predicted Score	Observed Score	Growth and p-value	
			Mean	Mean	Difference of Means	Count
Grade One	AERS Student	No	547.8	555.8	8.0	24
		Yes	547.4	527.9	-19.6	24
	Difference (Yes Minus No)		-0.4	-27.9	-27.5 (p-value=0.001)*	
Grade Two	AERS Student	No	568.7	580.0	11.3	17
		Yes	567.8	554.2	-13.6	17
	Difference (Yes Minus No)		-0.9	-25.8	-24.9 (p-value=0.004)*	
Grade Three	AERS Student	No	738.4	735.1	-3.4	76
		Yes	738.5	729.2	-9.3	76
	Difference (Yes Minus No)		0.0	-5.8	-5.9	
Grade Four	AERS Student	No	42.4	41.6	-.8	55
		Yes	42.2	37.5	-4.7	55
	Difference (Yes Minus No)		-0.2	-4.1	-3.9	
Grade Five	AERS Student	No	39.9	39.8	-.1	37
		Yes	39.7	33.4	-6.4	37
	Difference (Yes Minus No)		-0.2	-6.4	-6.2 (p-value=0.025)*	

*notates statistical significance

In every grade and in every type of pairing the predicted scores were within one point of each other indicating that the students were expected to perform in a similar manner, yet the observed means were always worse for the AERS students. For each test there were three grades whose growth difference mean was considered to be statistically significant, although it varied from grade to grade with each grade finding significance in two out of the three tests.

Conclusions and Considerations

The Additional Elementary Reading Support intervention exhibited mixed results. Overall, in four of the five grades, the mean student growth for the AERS students is smaller than the mean of the rest of the students in the particular grades. Yet, when broken down by school, there are grades where the AERS students performed better than the other students in their grade based upon the growth means.

We do not see any positive results for the AERS student growth when matching them with students who had similar predicted scores. This may be because once they are matched, the control student may also have been in an intervention; and if this intervention was in an AERS school, then the control student would probably have received his or her intervention from a certified teacher instead of from an educational aide.

There is some, but very little evidence to support the continuation of this particular program. A further study would require details on all students who receive intervention, the nature of the intervention, and the certification status of the intervention provider.

School Technology Challenge

In SY1314, Knox County Schools made a significant investment in computing devices and associated professional development to fully integrate technology within a subset of KCS schools. The resulting School Technology Challenge (STC) created a 1:1 student to device ratio with the main objective of increasing teacher effectiveness to drive increases in student outcome data. The STC theory of action highlights three areas of focus for meeting this main objective. The foci are listed below.

1) Increase the individualization and differentiation of student-centered instruction.
2) Increase student affinity, motivation and engagement in the classroom.
3) Increase the effectiveness of teaching through both 1) and 2) while integrating technology-based education aids.

Eleven schools were selected to participate in the STC via an application process. The participating schools are listed in the Table 4.1 below.

Table 4.1: STC Schools

School	Level
Bonny Kate Elementary	Elementary
Corryton Elementary	Elementary
Halls Elementary	Elementary
Mooreland Heights Elementary	Elementary
Norwood Elementary	Elementary
Sterchi Elementary	Elementary
Holston Middle	Middle
West Valley Middle	Middle
Vine Middle	Middle
Bearden High	High
South-Doyle High	High

It is important to note that this analysis constitutes a formative evaluation of the School Technology Challenge in Knox County. Another annual formative analysis will be completed at the end of the SY1415 to track progress in the second year of the initiative. A summative analysis is expected at the end of SY1617.

Methodology

Participating schools were selected through a non-random application process. This non-randomized assignment of treatment (the STC) requires the use of quasi-experimental methods. As a result, a pool of control schools serves as the counter-factual to the STC treatment.

Control schools were selected via partial least squares regression on SY1213 data. Fourteen independent variables were used to model multiple dependent variables (previous year

mean TCAP Normal Curve Equivalents, mean EOC or SAT-10 scaled scores and TVAAS effect sizes). The partial least squares regression created linear combinations of the 14 independent variables to best capture the variation in the dependent variables. The resulting models (one for each elementary schools, middle schools, and high schools) described anywhere from 55% (for elementary) to 77% (for high schools) of the total variance in the dependent variables.

Schools were grouped using hierarchal clustering based on the linear combinations of their independent variables. The nearest neighbor to a STC school was chosen as its control school match because these schools demonstrated similar outcomes from similar inputs in SY1213. This methodology for identifying control schools is analogous to a propensity score matching. The partial least squares method was used because we had no prior knowledge of the relative importance of each input variable on the output. See Table 4.2 below for a list of control schools.

Table 4.2: Control Schools

School	Level
Ball Camp Elementary	Elementary
Carter Elementary	Elementary
Copper Ridge Elementary	Elementary
Dogwood Elementary	Elementary
Powell Elementary	Elementary
Spring Hill Elementary	Elementary
Carter Middle	Middle
Karns Middle	Middle
Northwest Middle	Middle
Hardin Valley Academy	High
West High School	High

Readers should note that Vine Middle School and its control match have been removed from SY1314 analyses because of the limited and late device deployment that occurred at Vine Middle (January 2014 and only in the 8th grade). L&N STEM Academy is also not a part of this particular study because they have been implementing 1:1 device deployment since their inception.

Most of the STC analyses will consider results from the entire grade level pool of STC schools versus the entire grade level pool of control schools. Analysis of individual STC schools versus their best-matched control school is difficult due to variation in unobservable characteristics, lower statistical power, and higher uncertainty in the school level data.

The STC theory of action was intentionally broad and had potential impacts in many areas beyond the scope of traditional program evaluation. The positive or negative effects of the

STC in Knox County will therefore be determined through a preponderance of evidence rather than changes in a single metric. Accordingly, this first formative evaluation of the initiative is also intentionally broad. Data included in this analysis include teacher interview data, student survey data, attendance rates, classroom rubric scores, and standardized test results.

Results: Teacher Focus Groups – Common Themes

Principals at schools participating in the STC chose a subset of teachers to participate in small group interviews. In many cases, the common themes from these interviews verify results obtained through our qualitative studies.

Nearly universally, the teachers that were interviewed felt that students were more engaged in classwork when the technology was involved. Many interviewees perceived that participation was richer and deeper when conducted through Canvas. These teachers noted that students that typically were afraid to contribute in large group settings (notably less outgoing and Special Education students) were more willing to participate in moderated discussion via Canvas than in the traditional classroom setting. Similarly, the teachers that were interviewed appreciated the ability for students to work at their own pace, and do so discreetly. Students that were the furthest behind could be given augmented or additional assignments directly on Canvas without other students being aware.

Most of the interviewed groups noted that teaching was "different" and that their role as a teacher was evolving. Some teachers are creating a more student driven learning environment where more advanced students are taking the lead in teaching their classmates. Commonly, the interviewed teachers are relying on the students to solve low level technology issues that their fellow students are experiencing. Some of the interviewed teachers are also relying on their students to instruct the teachers themselves about how to best use the devices. Many have cited these actions as confidence and relationship building activities for their students.

However, the teachers participating in the focus groups were generally critical of the training and professional development regarding the STC initiative. They largely felt that the professional development that was offered was too theoretical. They hoped for more practical training about how to use the devices in the field. Strangely, very few of the interviewed teachers reported that they tried to collaborate with other KCS STC schools outside of any district level offerings.

Time was a continual theme in the group interviews. The teachers participating in the focus groups noted that they were investing non-trivial amounts of planning and off-contract time to transition to a technology-integrated classroom. Most of these teachers also noted some small, but potentially continual losses in instructional time (due to system log-ins and boot-ups, for example). Teachers interviewed at elementary schools noted that they had to create time to teach typing skills to younger students. In addition, some of the interviewed teachers were also dealing with internet access issues as the project came on-line (mostly in the secondary levels) which affected time on task. These same teachers generally noted that the

access issues were largely solved by the end of the first year of implementation. Anecdotally, losses seemed to have been most severe in early phases on the STC initiative. Most focus group participants became more comfortable with their work-around processes as the year progressed.

Continuity was another common concern. A majority of the participating elementary and middle school teachers were deeply concerned over what will happen to their students as they transition to a feeder school that is not participating in the STC. The interviewed high school teachers noted the difficult timing for device collection. Students had relied on the device to research and take notes all year but lost access to the device right before final exams. Teachers that were interviewed at high mobility schools also noted some difficulty bringing transfer students up to speed on the usage of the classroom technology, but on the whole did not find it an onerous burden.

A final common theme in the primary grades revolved around parent engagement in the technology challenge. Many assignments were completed and graded on Canvas. This resulted in less paperwork going home. Multiple schools reported that parents felt their students were doing less school work after implementation of the technology. Anecdotally, parents were finding it difficult to track their students' progress with the grading information in Canvas. Schools lead grass-roots communication efforts to keep parents informed of the volume and quality of school work being completed. At the end of last year, parents could not be granted access to Canvas due to software platform issues.

Despite these challenges, the teachers interviewed for the focus groups were unanimous in their support of the STC initiative. Almost all of the teachers that were interviewed felt that the integration of technology was worth it. Nobody expressed regret that their school had participated in STC. In fact, participating elementary teachers were near unanimous in their request for more devices to achieve greater densities in early grades.

Results: Technology Integration and Personalization Rubrics

The KCS Office of Innovation adapted two rubrics to measure the breadth and depth of technology implementation in the STC schools. The Technology Integration Matrix (TIM) measured the depth of device integration into the classroom in five domains; Active, Collaborative, Constructive, Authentic and Goal Directed. It was adopted from the Arizona K-12 Center for Professional Development. A separate rubric was created in-house to classify the levels of personalized learning occurring in the STC classrooms. The Personalized Leaning Environment (PLE) rubric covered four domains; Student Centered Instruction, Student Engagement; Assessment and Learning Environment. Data was collected from a random sample of 112 teachers at both control and STC schools at the end of SY1314. Aggregate level results are contained in the tables below. Maximum values in each row are shaded in orange.

Table 4.3: Technology Integration Matrix Results

School	Active					Collaboration					Constructive				
	Entry	Adoption	Adaptation	Infusion	Transformation	Entry	Adoption	Adaptation	Infusion	Transformation	Entry	Adoption	Adaptation	Infusion	Transformation
Tech - All	57.1%	24.1%	13.4%	4.5%	0.9%	62.5%	23.2%	14.3%			51.8%	25.0%	21.4%		1.8%
Elementary - All	64.1%	18.8%	10.9%	6.3%		65.6%	18.8%	15.6%			54.7%	17.2%	26.6%		1.6%
Bonny Kate	90.0%		10.0%			80.0%	10.0%	10.0%			80.0%		20.0%		
Corryton	71.4%	28.6%				78.6%	7.1%	14.3%			71.4%		28.6%		
Halls	90.0%			10.0%		90.0%	10.0%				60.0%	30.0%			10.0%
Mooreland Hts	12.5%	25.0%	37.5%	25.0%		12.5%	50.0%	37.5%			12.5%	37.5%	50.0%		
Norwood	88.9%	11.1%				77.8%	22.2%				77.8%	22.2%			
Sterchi	30.8%	38.5%	23.1%	7.7%		46.2%	23.1%	30.8%			23.1%	23.1%	53.8%		
Middle - All	36.0%	36.0%	28.0%			60.0%	24.0%	16.0%			36.0%	48.0%	16.0%		
Holston	40.0%	20.0%	40.0%			40.0%	40.0%	20.0%			40.0%	50.0%	10.0%		
West Valley	33.3%	46.7%	20.0%			73.3%	13.3%	13.3%			33.3%	46.7%	20.0%		
High - All	60.9%	26.1%	4.3%	4.3%	4.3%	56.5%	34.8%	8.7%			60.9%	21.7%	13.0%		4.3%
Bearden	60.0%	10.0%	10.0%	10.0%	10.0%	60.0%	20.0%	20.0%			70.0%		20.0%		10.0%
South-Doyle	61.5%	38.5%				53.8%	46.2%				53.8%	38.5%	7.7%		

Table 4.4: Technology Integration Matrix Results (Continued)

School	Authentic					Goal Directed				
	Entry	Adoption	Adaptation	Infusion	Transformation	Entry	Adoption	Adaptation	Infusion	Transformation
Tech - All	57.1%	32.1%	9.8%	0.9%		67.9%	25.0%	5.4%	1.8%	
Elementary - All	60.9%	31.3%	7.8%			71.9%	20.3%	7.8%		
Bonny Kate	80.0%	10.0%	10.0%			90.0%		10.0%		
Corryton	78.6%	14.3%	7.1%			85.7%	7.1%	7.1%		
Halls	60.0%	40.0%				80.0%	10.0%	10.0%		
Mooreland Hts	37.5%	50.0%	12.5%			25.0%	75.0%			
Norwood	66.7%	33.3%				88.9%	11.1%			
Sterchi	38.5%	46.2%	15.4%			53.8%	30.8%	15.4%		
Middle - All	48.0%	32.0%	20.0%			60.0%	36.0%		4.0%	
Holston	40.0%	20.0%	40.0%			40.0%	50.0%		10.0%	
West Valley	53.3%	40.0%	6.7%			73.3%	26.7%			
High - All	56.5%	34.8%	4.3%	4.3%		65.2%	26.1%	4.3%	4.3%	
Bearden	60.0%	20.0%	10.0%	10.0%		60.0%	30.0%		10.0%	
South-Doyle	53.8%	46.2%				69.2%	23.1%	7.7%		

Table 4.5: PLE Rubric Results

School	Student Centered Instruction				Student Engagement			
	Not Evident	Emerging	Evident	Exemplary	Not Evident	Emerging	Evident	Exemplary
Tech-All	30.0%	50.2%	16.4%	3.4%	10.3%	48.7%	36.6%	4.5%
Elementary-All	32.5%	49.7%	14.4%	3.4%	8.6%	44.5%	43.8%	3.1%
Bonny Kate	50.0%	44.0%	6.0%	0.0%	15.0%	30.0%	55.0%	0.0%
Corryton	54.3%	35.7%	10.0%	0.0%	21.4%	50.0%	28.6%	0.0%
Halls	42.0%	44.0%	10.0%	4.0%	5.0%	55.0%	35.0%	5.0%
Mooreland Hts	15.0%	57.5%	22.5%	5.0%	0.0%	43.8%	43.8%	12.5%
Norwood	8.9%	71.1%	20.0%	0.0%	5.6%	61.1%	33.3%	0.0%
Sterchi	15.4%	53.8%	20.0%	10.8%	0.0%	30.8%	65.4%	3.8%
Middle-All	18.4%	60.0%	21.6%	0.0%	6.0%	64.0%	26.0%	4.0%
Holston	8.0%	74.0%	18.0%	0.0%	0.0%	90.0%	10.0%	0.0%
West Valley	25.3%	50.7%	24.0%	0.0%	10.0%	46.7%	36.7%	6.7%
High-All	35.7%	40.9%	16.5%	7.0%	19.6%	43.5%	28.3%	8.7%
Bearden	38.0%	32.0%	14.0%	16.0%	15.0%	35.0%	30.0%	20.0%
South-Doyle	33.8%	47.7%	18.5%	0.0%	23.1%	50.0%	26.9%	0.0%

Table 4.6: PLE Rubric Results (Continued)

School	Assessment				Learning Environment			
	Not Evident	Emerging	Evident	Exemplary	Not Evident	Emerging	Evident	Exemplary
Tech-All	44.3%	40.2%	15.2%	0.3%	20.3%	36.2%	37.9%	5.6%
Elementary-All	50.0%	37.6%	11.3%	1.0%	16.4%	33.6%	44.5%	5.5%
Bonny Kate	96.7%	3.3%	0.0%	0.0%	27.5%	42.5%	30.0%	0.0%
Corryton	62.8%	27.9%	7.0%	2.3%	32.1%	35.7%	32.1%	0.0%
Halls	60.0%	33.3%	6.7%	0.0%	20.0%	10.0%	67.5%	2.5%
Mooreland Hts	33.3%	54.2%	12.5%	0.0%	3.1%	46.9%	43.8%	6.3%
Norwood	29.6%	40.7%	29.6%	0.0%	5.6%	38.9%	44.4%	11.1%
Sterchi	17.1%	63.4%	14.6%	4.9%	3.8%	30.8%	51.9%	13.5%
Middle-All	25.3%	49.3%	25.3%	0.0%	24.0%	50.0%	24.0%	2.0%
Holston	3.3%	76.7%	20.0%	0.0%	12.5%	80.0%	7.5%	0.0%
West Valley	40.0%	31.1%	28.9%	0.0%	31.7%	30.0%	35.0%	3.3%
High-All	47.8%	36.2%	14.5%	1.4%	27.2%	28.3%	34.8%	9.8%
Bearden	43.3%	36.7%	16.7%	3.3%	25.0%	25.0%	27.5%	22.5%
South-Doyle	51.3%	35.9%	12.8%	0.0%	28.8%	30.8%	40.4%	0.0%

It is notable that the majority of teachers observed on the TIM rubric were at the earliest stage of development for technology integration. This may be a reflection of the experimental nature of the first year of STC implementation.

The PLE results show similar trends. The most frequently observed results for the Learning Environment domain is in the "Evident" category. However, the bulk of the questions in the Learning Environment domain are based on the physical set-up of the classrooms. When we examine the Learning Environment subdomain that examines fundamental shifts in how content is delivered, we find that the most frequent response is in the "Emerging" category.

A statistical comparison between rubric scores at STC and control schools is contained below. The qualitative scores from the rubrics were translated to a quantitative scale (1="Not Evident", 2="Emerging", etc.). Previous studies conducted by the REA team have indicated growth in TEAM scores in a teacher's first five years of employment. Therefore, only veteran teachers with more than five years of experience are considered in this portion of the analysis to prevent any bias.

Table 4.7: Mean Rubric Statistics

Measure	Type	N	Mean Rubric Score	Std. Deviation	p-value
AVG TIM	Tech	88	1.55	0.62	0.000*
	Control	48	1.05	0.17	
AVG PLE	Tech	88	2.03	0.49	0.000*
	Control	48	1.45	0.30	

*notates statistical significance

The null hypothesis, that the mean TIM and PLE scores are no different between STC and control schools can be rejected. Among veteran teachers, the mean TIM score was 1.55 suggesting that the average veteran STC teacher scored between the "Entry" and "Adoption" levels of technology integration. The mean PLE score for this same subset of teachers was 2.03, suggesting that the average veteran teacher was categorized as slightly above the "Emerging" level for personalization of the learning experience.

These results mesh with the information obtained in the teacher interviews. Although there have been some changes to the way content is delivered in STC schools, it has not been large enough to fundamentally change the classroom experience. In fact, the only STC school that mentioned deploying truly blended learning environments during the teacher interviews was Mooreland Heights Elementary (and there only really in 1st grade).

The analysis that follows this data should be considered within the context of the survey and interview data. In essence, the first year of implementation of the STC was a year of trial, error and experimentation. The level of personalization and technology integration may not yet be deep enough to impact the other indicators of success that are presented in this study. This

hypothesis is based on the sample of PLE and TIM rubric scores and corroborated by teacher interview data. Readers are reminded that this is a formative analysis of a long term study.

Results: Student Survey Data

The TRIPOD survey data was collected for individual teachers at STC and control schools. TRIPOD is a research-based school climate survey so it is not directly related to technology integration. Baseline data (domain normal curve equivalents) was from SY1213. The change in NCE (in each domain) was calculated for each teacher as the NCE in SY1314 minus the NCE in SY1213. The results were then aggregated by STC designation (STC or control). The STC theory of action targeted changes in the Cares, Captivates, Confers, and Controls domains. Results are contained below.

Table 4.8: TRIPOD Data

		All Teachers				Veteran Teachers			
Domain	**School Type**	**N**	**Mean Δ NCE**	**Std. Deviation**	**p-value**	**N**	**Mean Δ NCE**	**Std. Deviation**	**p-value**
Cares	Tech	119	-1.2	22.9	0.379	97	-1.4	22.2	0.319
	Control	136	-3.6	20.3		105	-4.3	20.4	
Challenge	Tech	119	-3.2	21.2	0.628	97	-3.6	21.4	0.644
	Control	136	-1.9	19.0		105	-2.3	19.2	
Captivates	Tech	119	-3.1	22.7	0.108	97	-3.5	22.7	0.228
	Control	136	-7.5	20.6		105	-7.3	21.2	
Confers	Tech	119	4.9	23.0	0.345	97	5.1	22.8	0.514
	Control	136	2.3	20.2		105	3.1	20.4	
Consolidates	Tech	119	0.0	21.7	0.321	97	-0.5	21.2	0.432
	Control	136	-2.5	18.3		105	-2.7	18.3	
Controls	Tech	119	1.9	24.7	0.601	97	1.5	23.5	0.467
	Control	136	0.3	24.4		105	-1.0	24.5	
Clarifies	Tech	119	0.8	29.2	0.261	97	0.0	28.9	0.829
	Control	136	-3.1	25.8		105	-0.8	26.3	

Directionally, the mean change in NCE was more favorable for STC schools in the targeted domains. In fact, the mean change in NCE was more favorable for STC schools in every domain except for Challenge. However, the results are only directional. The p-value reported in Table 4.8 above is the probability that the difference in mean NCE occurs by chance. The most interesting result is that the probability of the difference in change in Captivates NCE is due to chance is only 11% (when using the entire pool of teachers, 23% when using the veteran teacher pool). This information confirms other student level survey data collected regarding only the STC initiative and will be discussed later in this analysis

Student perception data collected specifically around the STC are listed below. It was common response at all grade levels that there was more interest in classes when the technology was

deployed. It was also a common response that respondents felt more responsible for their own work. This information mirrors teacher perceptions that were collected during focus groups.

Table 4.9: Student Survey Data

Question	Elementary Responses (N=476)		Middle School Responses (N=1561)		High School Responses (N=1593)	
	No	Yes	No	Yes	No	Yes
Is it easier for you to complete schoolwork on your own (do research, write papers)?	17.4%	82.6%	10.7%	89.3%	9.1%	90.9%
Do you like being able to communicate with your teacher by email / messages?	33.6%	66.4%	11.9%	88.1%	10.9%	89.1%
Do you get your grades or scores more quickly?	22.4%	77.6%	12.8%	87.2%	22.5%	77.5%
Are you more interested in class when you use your computer?	13.5%	86.5%	16.2%	83.8%	20.9%	79.1%
Are you asked to do more projects?	43.9%	56.1%	25.1%	74.9%	25.8%	74.2%
Do you work in groups more often?	35.9%	64.1%	38.8%	61.2%	39.2%	60.8%
Do you find it easy to understand the directions your teacher gives you about using the computer?	14.3%	85.7%	13.8%	86.2%	13.1%	86.9%
Do you sometimes find yourself distracted in class because of your computer?	76.1%	23.9%	65.1%	34.9%	54.3%	45.7%
Are you worried that you will lose or break the computer?	66.7%	33.3%	56.7%	43.3%	58.4%	41.6%
Do you feel like your schoolwork has gotten more difficult since you got your computer?	80.7%	19.3%	68.3%	31.7%	67.8%	32.2%
Do you feel more responsible for your own schoolwork?	11.3%	88.7%	14.0%	86.0%	16.2%	83.8%

There are other trends in the data that are worth noting. Teachers interviewed through the focus groups noted that some students were distracted with the technology. However, they felt that the general level of distraction in their classroom was the same pre and post deployment of the technology. In essence, teachers said that the technology was a source of distraction, but did not perceptively increase the level of distraction from an ordinary classroom setting. The responses in the student perception survey were more mixed than the teacher perceptions. There is a notable increase in respondents who felt they were distracted as the grade level of the student increases. There is also an interesting discontinuity in the data regarding project work. Elementary respondents did not perceive nearly the change in project work as the secondary students.

Please note that it is possible that these trends have nothing to do with the integration of technology. General distraction may rise in classrooms as students grow older with or without technology. Also, elementary students may have been doing more project work to begin with, so

did not notice an increase with the focus on project based learning. We will continue to monitor trends in the data and perhaps collect some control school data to better make sense of the trends.

The most notable result in the above survey data is just how few of respondents felt school work had gotten more difficult after the deployment of the devices. The survey responses seem to echo the TRIPOD data in the challenges domain.

One goal of the STC initiative was to increase students' familiarization with technology. The hypothesis is that foundational skills in technology will support later college and career readiness. Survey data (pre-deployment) was collected by the Parthenon Group as part of the KCS smarter spending initiative. REA conducted a post survey near the end of SY1314. The changes in responses are noted in Table 4.10 below.

Table 4.10: Student Survey Data - Activities

Task	Percent of Students who Report Never Having Done the Following Computer Driven Tasks								
	Elementary			Middle			High		
	Pre (N=799)	Post (N=476)	Change	Pre (N=1741)	Post (N=1561)	Change	Pre (N=2387)	Post (N=1539)	Change
Typing an essay	38.7%	10.0%	-28.7%	8.7%	2.9%	-5.8%	12.7%	3.6%	-9.1%
Using the internet to do research	19.4%	4.6%	-14.9%	3.7%	0.5%	-3.2%	4.0%	0.5%	-3.5%
Creating a multimedia presentation (video, slides, or both)	53.8%	12.9%	-40.9%	12.6%	1.9%	-10.8%	17.6%	2.0%	-15.5%
Using the internet to communicate with others (Email, blogs)	58.1%	55.3%	-2.9%	35.0%	25.3%	-9.8%	21.4%	9.0%	-12.4%
Creating graphs and charts using data	60.2%	24.2%	-35.9%	42.7%	15.2%	-27.5%	40.7%	23.2%	-17.6%
Taking quizzes or tests	4.1%	1.5%	-2.6%	5.2%	0.8%	-4.5%	10.3%	1.6%	-8.7%
Practice math or reading on a program or app	8.5%	8.8%	0.3%	34.6%	17.8%	-16.8%	51.8%	17.2%	-34.6%
Playing educational games	9.9%	2.9%	-7.0%	16.5%	6.3%	-10.3%	38.3%	27.9%	-10.4%
Using Canvas for classwork	38.9%	2.4%	-36.5%	2.8%	34.5%	31.7%	7.0%	1.2%	-5.8%

The increase in responses seems to indicate that students are more familiar with the tasks in the survey. Trends were strong across all grade levels. This is some evidence that the technology challenge is meeting its intended effect of increasing student familiarity with technology.

Results: Student Attendance

The STC theory of action was hypothesized to lead to increases in student attendance. It was anticipated that increases in student engagement and affinity would be reflected in increased student attendance rates. Attendance data was normalized by considering a change in attendance rather than the number of days attended. The change in attendance was calculated as each student's SY1314 attendance rate (as percent of possible days) minus the student's SY1213 attendance rate. Data are aggregated by STC status.

Table 4.11: Attendance

Change in Attendance	N	Mean Δ Attend	Std. Deviation	p-value
Tech	5701	-0.40%	5.54%	0.14
Control	6469	-0.25%	5.54%	

With a p value of 0.14, we fail to reject the null hypothesis that there is no difference in the change in attendance rate between the STC and control schools. Directionally, both STC and control schools experienced a decrease in mean attendance. The magnitude of the change in the STC schools was slightly higher.

We further disaggregate the data into two specific groups. The first group is made up of students who attended 90% or more instructional days in SY1213. Previous modeling work by REA has indicated this attendance threshold is a predictor of future success. There is no statistical difference between the mean change in attendance rate between STC and control schools among this subset of students.

Table 4.12: Aggregate Attendance Data: Students with Greater than 90% Attendance

Change in Attendance	N	Mean Δ Attend	Std. Deviation	p-value
Tech	4994	-0.48%	3.76%	0.11
Control	5665	-0.37%	3.71%	

The analysis is repeated for students who missed more than 10% of instructional days in SY1213. Again, there is no statistical difference between mean change in attendance rates between STC and control schools.

Table 4.13: Aggregate Attendance Data: Students with Less than 90% Attendance

Change in Attendance	N	Mean Δ Attend	Std. Deviation	p value
Tech	707	0.16%	12.15%	0.53
Control	804	0.55%	12.23%	

It is interesting to note the directionality in the data. Among each population subset, the STC schools had less favorable mean changes in attendance rates.

Results: Observation Data

The STC theory of action is intended to affect five indicators scored on the NIET instruction rubric. These five indicators are Teacher Knowledge of Students (TKS), grouping (GRP), Academic Feedback (FEED), Presenting Instructional Content (PIC) and Activities and Materials (ACT). The mean change in indicator score was calculated using the following procedure. For each teacher, the mean indicator score for SY1314 was subtracted from the mean indicator score for SY 1213. The mean of these differences is contained below (for veteran teachers only).

Table 4.14: STC Teacher Observation Scores: Indicator Level

Indicator	School Type	N	Mean Change in Score	Std. Deviation	p-value
ACT	Tech	355	-0.075	0.599	0.077
	Control	373	0.005	0.627	
FEED	Tech	355	0.047	0.611	0.398
	Control	373	0.086	0.642	
GRP	Tech	355	-0.032	0.567	0.008*
	Control	373	0.083	0.598	
PIC	Tech	355	-0.056	0.665	0.088
	Control	373	0.029	0.663	

The probability that differences in the change in the Grouping indicator score occurs by chance is 0.8%. The mean change in GRP score is therefore statistically higher in control schools (at the 95% confidence limit). The difference in mean change in ACT and PIC are statistically significant at the 90% confidence limit. The change in mean indicator score is higher for control schools compared to their STC counterparts. They data may be a reflection of teachers adjusting their practice to accommodate the technology. Participants in the teacher focus groups mentioned impacts on lesson planning and taking a more trial-and-error approach to teaching than they have in the past.

It is also possible that the data is too variable to contribute to the analysis. The observation data is difficult to analyze due to changes in the observation structure and competing systems. In SY1314, teachers could choose if their final classroom visit was announced or unannounced. This makes year to year comparisons difficult. In addition, there are two schools that use the TAP observation system in the pool of STC schools. An attempt at balance was made by including two TAP schools in the control group, but the number of TAP teachers in the STC pool is much larger than the corresponding number of TAP teachers in

the control pool. For these reasons, observation indicator scores are not recommended as key measures in the STC analysis going forward.

Results: Student Outcomes – TVAAS Data

The overarching goal of the STC was to impact student outcome data through a deep level of individualization. Results of t-tests on the one year mean TVAAS index are contained below. TVAAS indices could be used in the analysis because the effect of school size on TVAAS index was controlled for in the creation of the control schools.

Table 4.15: School-wide TVAAS

Subject Area	School Type	N	Mean Index	Std. Deviation	p-value
Literacy	Tech	10	-0.7	1.83	0.343
	Control	10	-1.7	2.67	
Numeracy	Tech	10	0.6	4.30	0.589
	Control	10	-0.4	3.81	
Literacy and Numeracy	Tech	10	-0.4	3.50	0.537
	Control	10	-1.4	3.60	
Overall	Tech	10	0	4.16	0.723
	Control	10	-0.8	5.65	

Directionally, the mean (school-level) TVAAS indices were higher in the STC schools compared to the control schools. However, the results are not statistically significant at typical confidence limits.

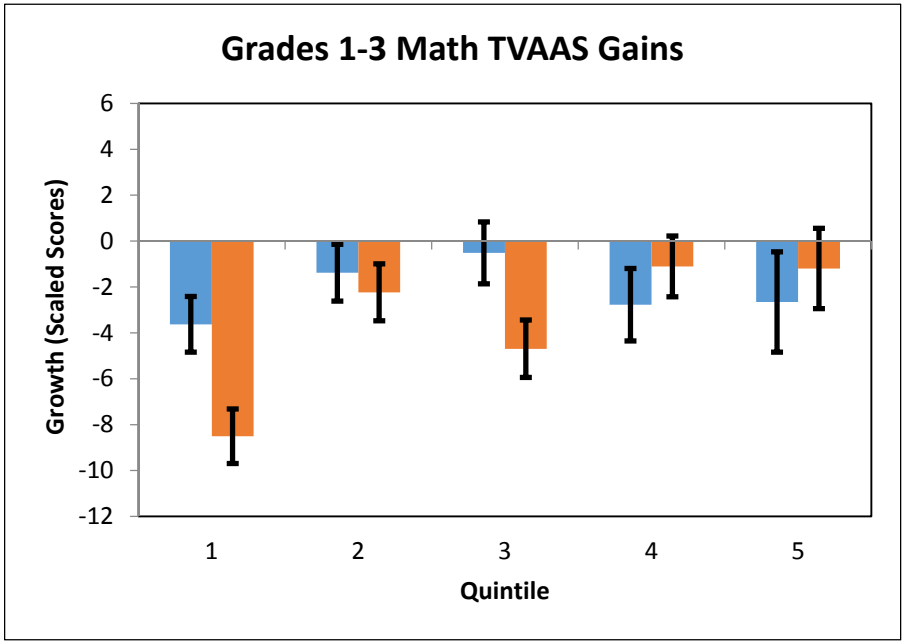
The same general trend appears in the teacher level TVAAS gains. Results are contained below.

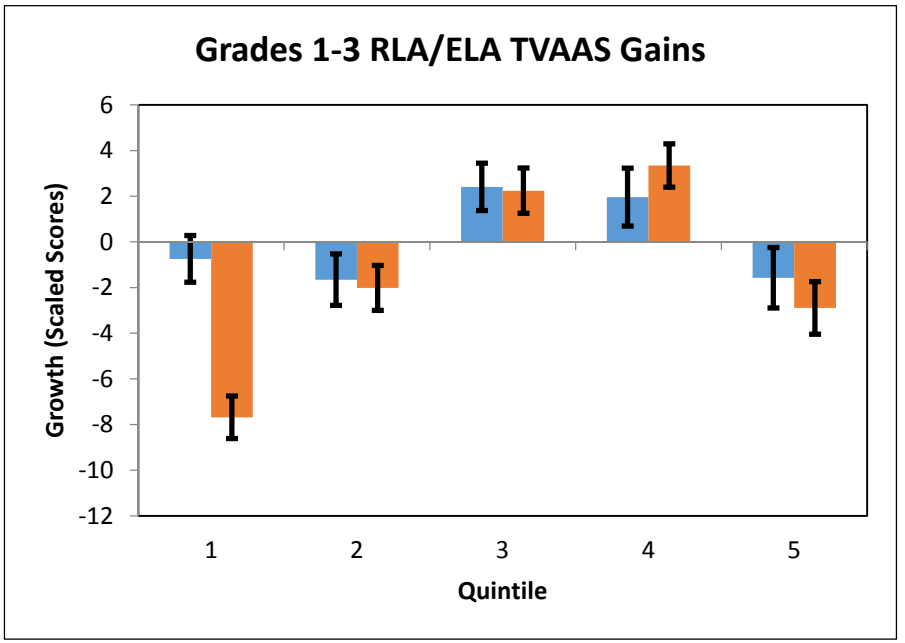
Table 4.16: Grade Band TVAAS

Level	School Type	N	Mean TVAAS Gain	Std. Deviation	p-value
Elementary - Grades 1-3	Tech	229	-0.86	7.35	0.597
	Control	309	-1.21	7.60	
Elementary - Grades 4-5	Tech	145	-1.59	4.27	0.251
	Control	192	-1.08	3.89	
Middle	Tech	97	1.57	4.64	0.467
	Control	107	1.02	5.96	
High	Tech	96	1.98	10.39	0.832
	Control	85	1.65	10.55	

Directionally, the mean TVAAS gain was higher for STC schools in all grade bands except grades 4-5. The differences, however, were not statistically significant (at the 95% confidence limit).

It was hypothesized that the STC would drive greater gains in lower performing students and aid in closing performance gaps because of deeper individualization and personalization. To this end, school-level TVAAS gains were analyzed by performance quintile (1 is the lowest and 5 is the highest). Higher growth in the lower quintiles would be evidence of accelerated learning among the students that were furthest behind. Results by grade level are contained below.

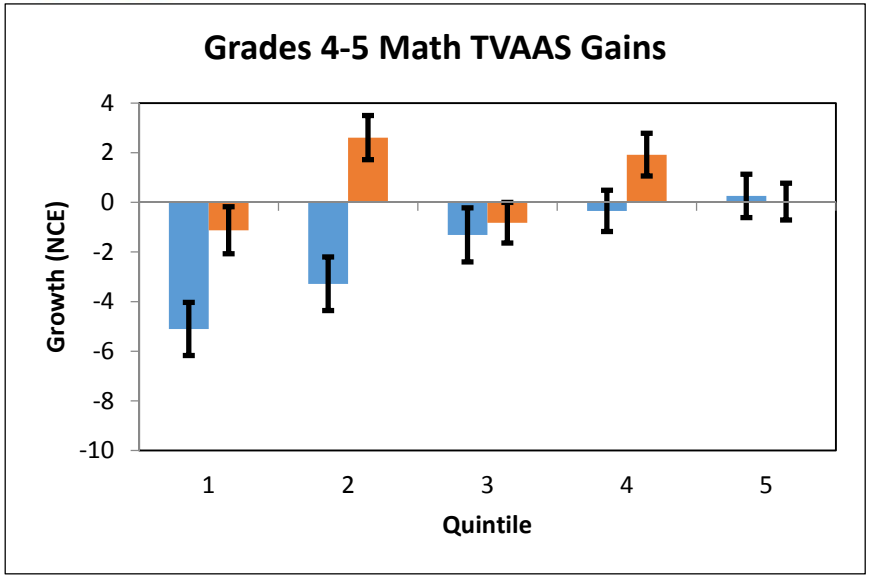


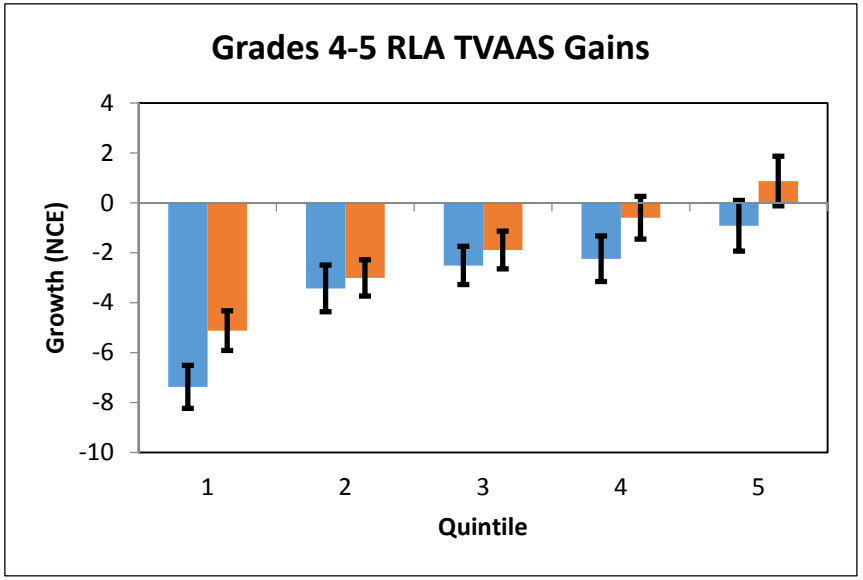


		Quintile				
		1 (Lowest)	*2*	*3*	*4*	*5 (Highest)*
Control	Avg Growth	-8.51	-2.24	-4.69	-1.10	-1.20
	Std Error	1.19	1.24	1.25	1.33	1.75
Tech	Avg Growth	-3.63	-1.38	-0.52	-2.78	-2.66
	Std Error	1.21	1.24	1.35	1.58	2.19

		Quintile				
		1 (Lowest)	*2*	*3*	*4*	*5 (Highest)*
Control	Avg Growth	-7.69	-2.02	2.24	3.34	-2.89
	Std Error	0.93	0.99	0.99	0.95	1.15
Tech	Avg Growth	-0.75	-1.66	2.40	1.96	-1.57
	Std Error	1.02	1.13	1.04	1.27	1.33

Figure 4.1: TVAAS Distributions by Quintile: Grades 1-3

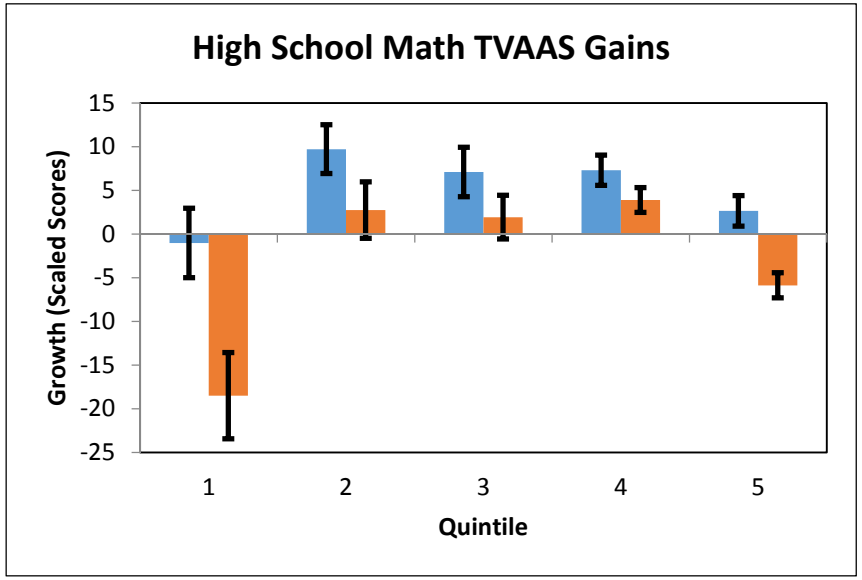


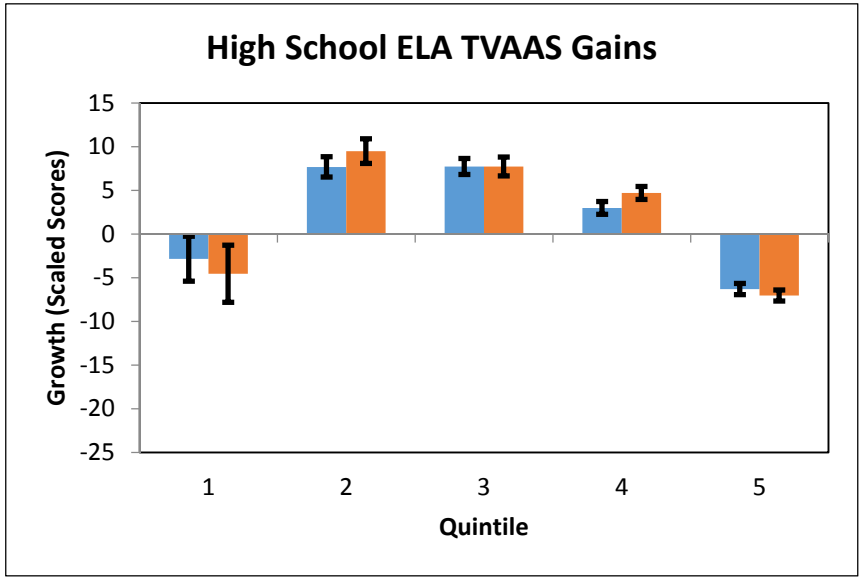


		Quintile				
		1 (Lowest)	***2***	***3***	***4***	***5 (Highest)***
Control	Avg Growth	-1.13	2.61	-0.82	1.92	0.03
	Std Error	0.95	0.89	0.82	0.86	0.74
Tech	Avg Growth	-5.11	-3.29	-1.31	-0.35	0.25
	Std Error	1.07	1.08	1.09	0.83	0.87

		Quintile				
		1 (Lowest)	***2***	***3***	***4***	***5 (Highest)***
Control	Avg Growth	-5.12	-3.01	-1.89	-0.60	0.87
	Std Error	0.79	0.73	0.76	0.86	0.99
Tech	Avg Growth	-7.37	-3.43	-2.51	-2.24	-0.92
	Std Error	0.86	0.93	0.76	0.91	1.02

Figure 4.2: TVAAS Distributions by Quintile: Grades 4-5





		Quintile				
		1 (Lowest)	***2***	***3***	***4***	***5 (Highest)***
Control	Avg Growth	-18.50	2.74	1.93	3.90	-5.87
	Std Error	4.94	3.23	2.50	1.42	1.44
Tech	Avg Growth	-1.03	9.72	7.10	7.30	2.65
	Std Error	3.98	2.80	2.83	1.73	1.76

		Quintile				
		1 (Lowest)	***2***	***3***	***4***	***5 (Highest)***
Control	Avg Growth	-4.55	9.49	7.73	4.71	-7.04
	Std Error	3.26	1.41	1.09	0.74	0.63
Tech	Avg Growth	-2.83	7.68	7.73	2.99	-6.30
	Std Error	2.57	1.16	0.91	0.73	0.64

Figure 4.3: TVAAS Distributions by Quintile: High School

Visual inspection of the plots indicated that there appears to be no systematic boost to TVAAS gains in STC schools. There are promising results in grades 1-3 at the lowest quintile, but these differences may be due to chance alone. With a few exceptions (notably Mooreland Heights Elementary), the early grades were the groups of students with the lowest technology densities. Also, any effect in the first quintile did not seem to translate to the second quintile, which leaves some room for doubt that the STC was the driving force behind the gains in the first quintile.

The high school math data also exhibits an interesting trend. This is the only subject/grade combination in which STC schools exhibit a significant and systematic increase in gains (when compared to the control schools). Interestingly, one of the STC high schools felt that integration of technology in math subjects was the most problematic. It is possible that the differences in the math gains are related to some other phenomenon (teacher turn-over, etc.) However, the fact that the growth in each STC quintile is at or above the standard error bar for the control school makes it a bit of an outlier when compared to other growth data. We will continue to monitor this data in the future to determine if the trend continues.

An analysis was done to determine how TVAAS index varies with the ratings on the PLE and TIM rubrics. The TVAAS index was chosen as the dependent variable for this analysis because indices can be compared between the different TVAAS models. The data cannot be broken down by grade level bands because of the sample sizes involved. However, there may be bias with using TVAAS index, because there are trends with class size and TVAAS index, but screening models indicate that class size accounts for only 5% of the total variation in TVAAS index.

Veteran teachers at STC schools were categorized according to how they were scored on the PLE rubric. Teachers in the bottom quintile for PLE scores were placed in the lower group. Teachers in the upper quintile were placed in the upper quintile. The resulting sample sizes may not be equal due to ties in the scoring at the cut-off point.

Table 4.17: TVAAS by PLE Quintile

PLE Quintile	N	Mean TVAAS Index	Std. Deviation	P value
Lower	10	0.21	1.60	.493
Upper	22	-0.52	3.09	

Directionally, the mean TVAAS index for teachers that scored in the upper quintile on the PLE rubric was lower than the mean TVAAS index for teachers who scored in the lower quintile on the PLE rubric. There was no statistical difference between mean TVAAS index

at typical confidence limits. It should be noted that low N counts decrease the statistical power of this test.

The same procedure was repeated regarding the TIM rubric. Teachers were categorized according to their TIM quintiles. Teachers in the lower quintile had lower observed TIM scores.

Table 4.18: TVAAS by TIM Quintile

TIM Quintile	N	Mean TVAAS Index	Std. Deviation	P value
Lower	9	0.23	1.72	.172
Upper	16	-1.26	2.87	

Directionally, the mean TVAAS index was again higher for teacher who scored in the lower quintile on the TIM rubric, without statistical significance at typical confidence limits.

Linear models were built to try to find a relationship between TVAAS index and individual indicator scores on the rubrics. Only the "Authentic" indicator on the TIM rubric had a statistically significant relationship with TVAAS index (at the 95% confidence limit). The Authentic indicator is scored regarding how "students use technology tools to solve real world problems meaningful to them, such as digital citizenship." The resulting linear model between the "Authentic" indicator score and the TVAAS index indicates that as the "Authentic" score increases, the TVAAS index decreases. The slope of the linear model indicates that for every point of increase in the "Authentic" indicator, the TVAAS index decreases by 1.66 points.

Table 4.19: Authentic Indicator Linear Model

Model Summary (N=32)			
p-value	R-squared	Slope	Intercept
0.021	0.165	-1.66	2.35

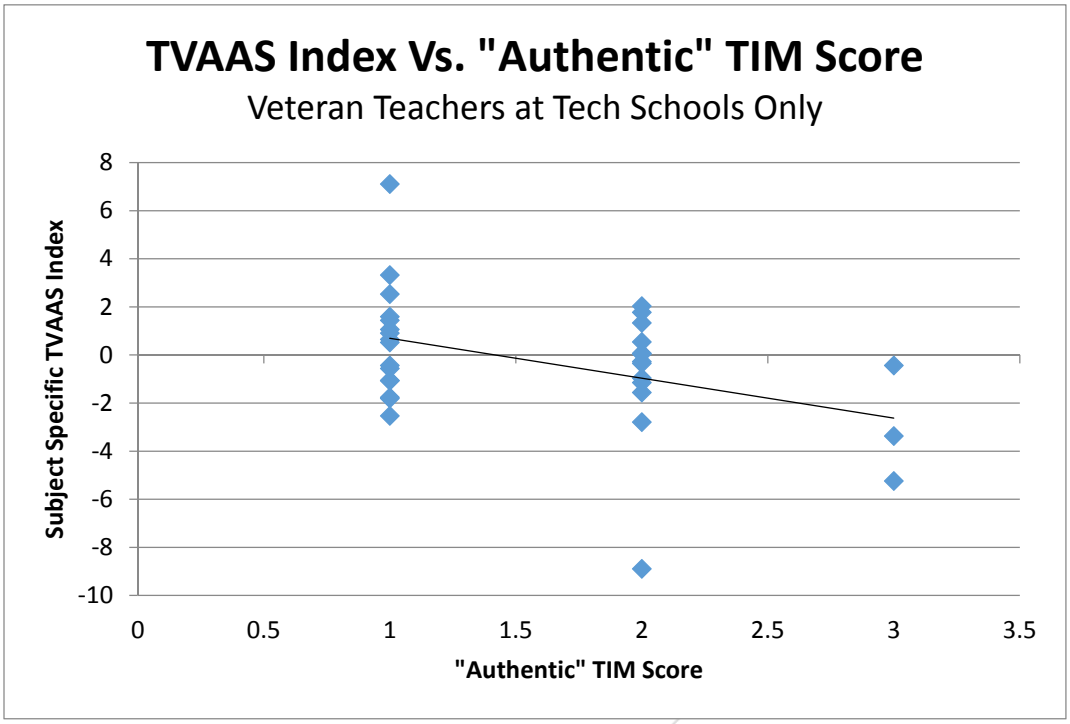


Figure 4.4: TVAAS Index vs. Authentic TIM Score

There are no statistically significant relationships between PLE indicator scores and teacher level TVAAS index.

Results: Student Outcomes – Achievement Data

A "difference in differences" model is used to analyze the proficiency level data from the SY1314 EOC and TCAP assessments. The "difference in differences" model uses the change in the data at the control schools as the baseline for changes in proficiency levels. Results are grouped by subject and grade bands in Figures 4.4 – 4.11 below. Favorable results at STC schools would cause the gray bar to be above the 0% line for the change in the percent of proficient or advanced students (% P/A) and the change in the percent of advanced students (% advanced). Favorable results for STC schools would have the gray bar below the 0% line for the percent of students below basic (% Below Basic).

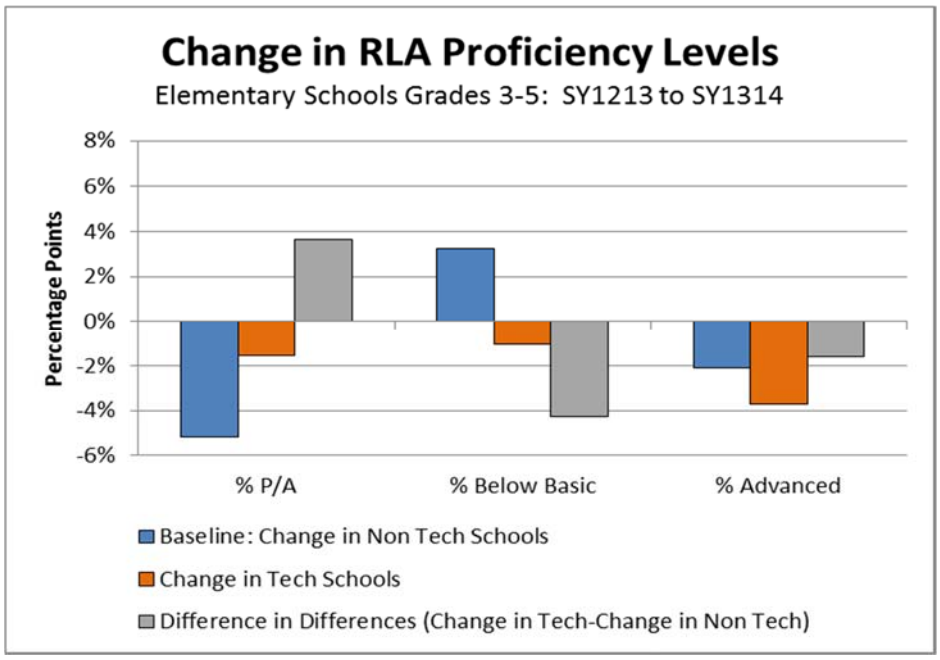


Figure 4.4: Change in RLA Proficiency: Elementary

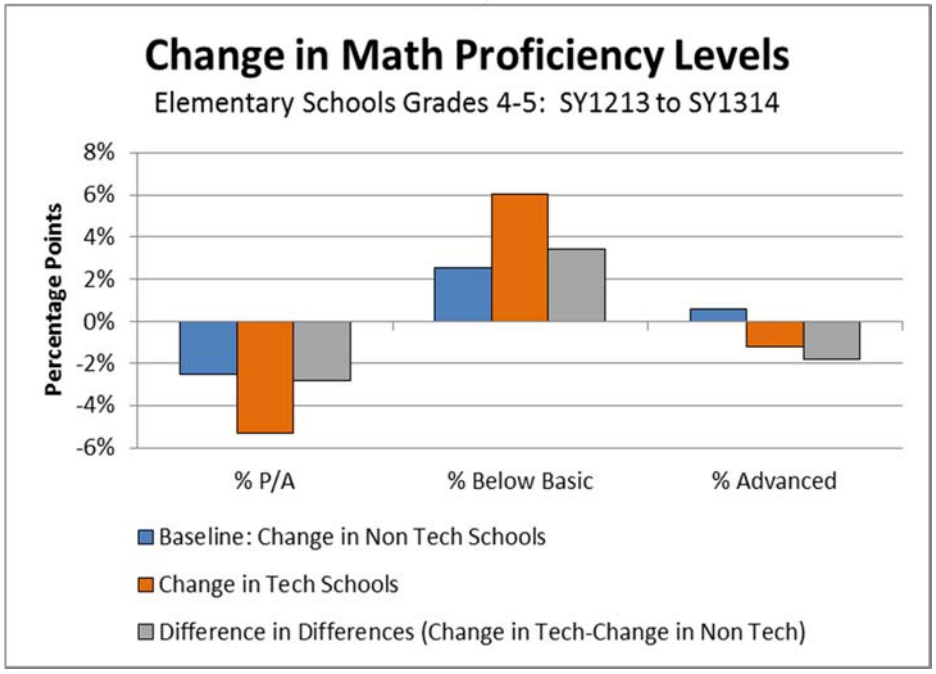


Figure 4.5: Change in Math Proficiency: Elementary

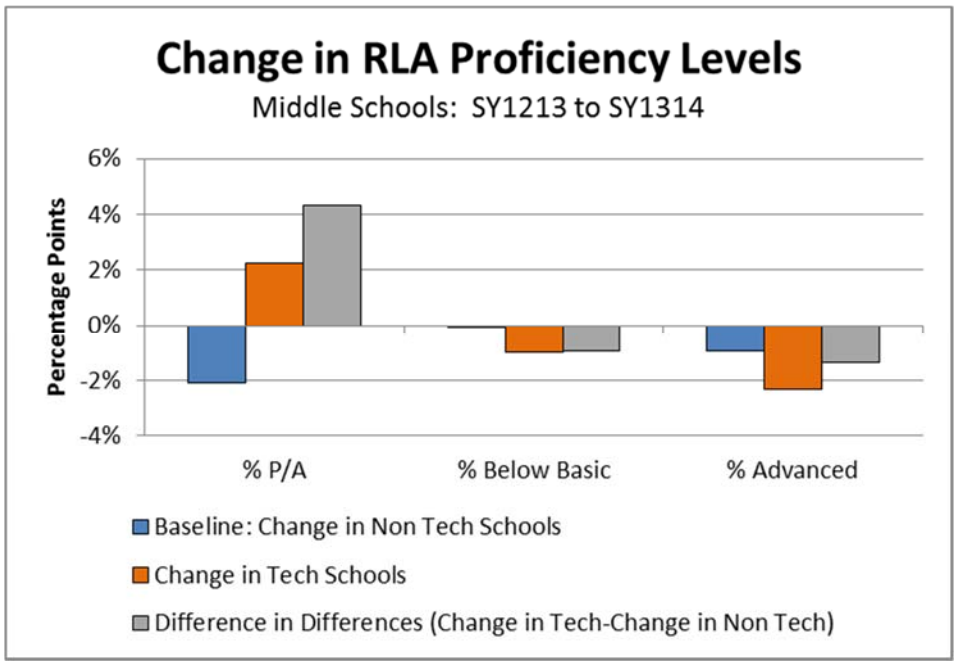


Figure 4.6: Change in RLA Proficiency: Middle

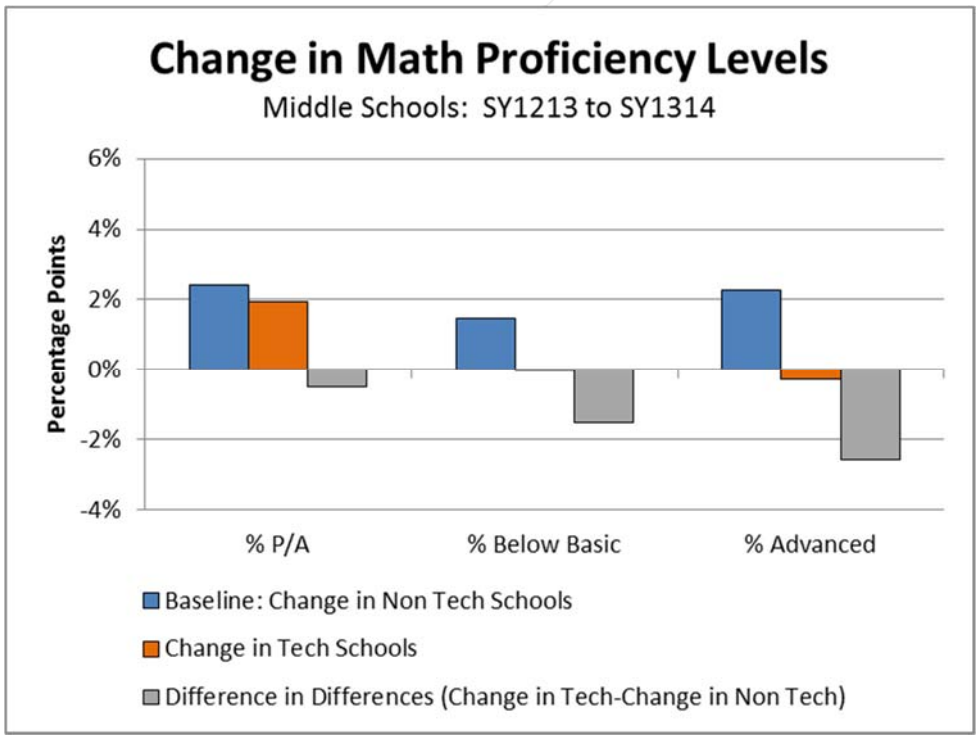


Figure 4.7: Change in Math Proficiency: Middle

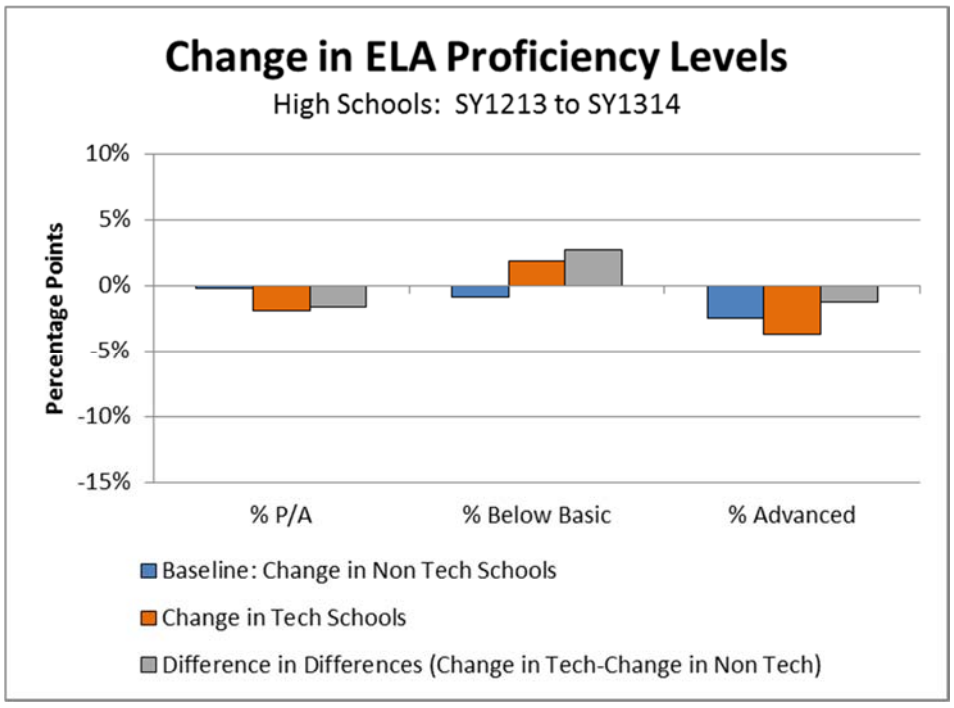


Figure 4.8: Change in ELA Proficiency: High

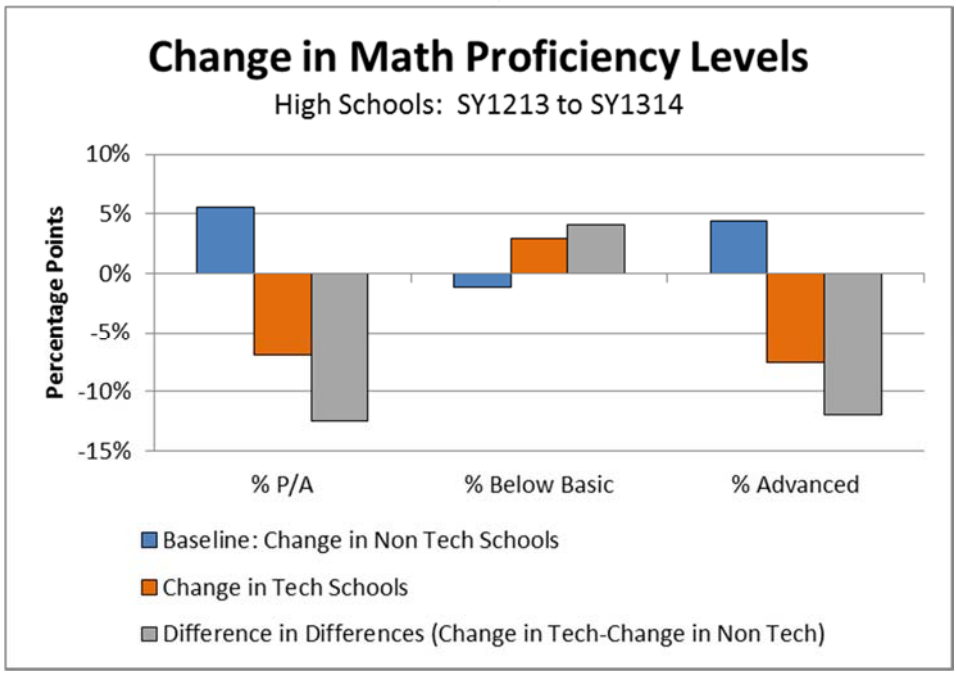


Figure 4.9: Change in Math Proficiency: High

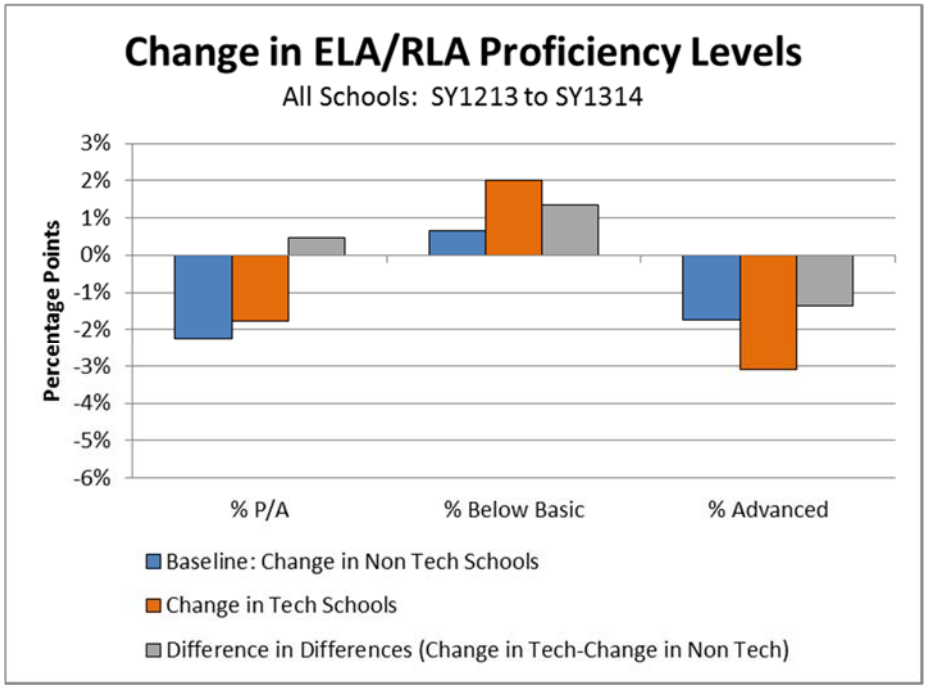


Figure 4.10: Change in ELA/RLA Proficiency: All Schools

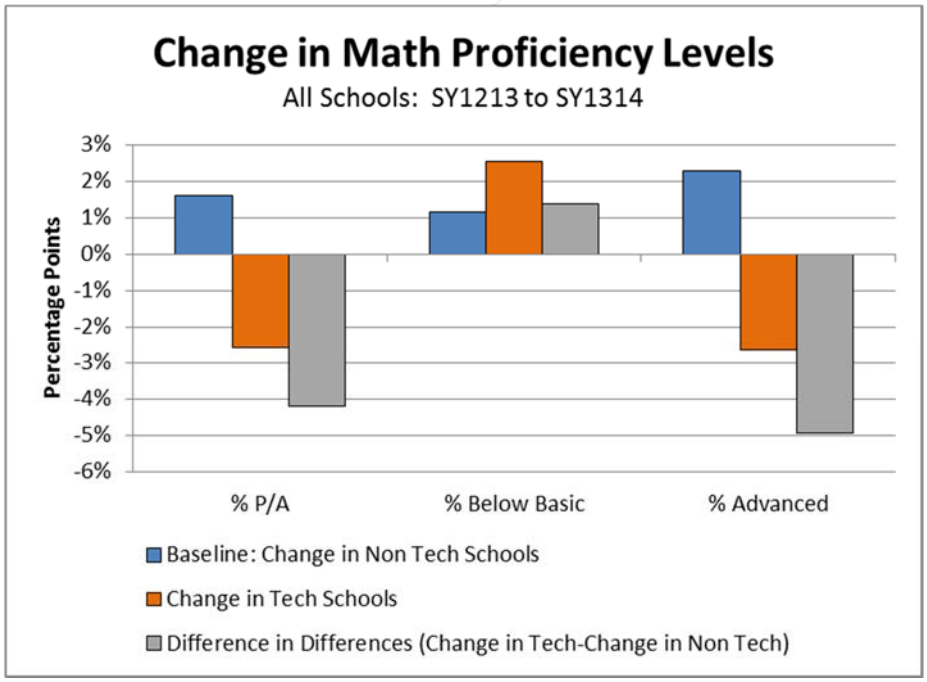


Figure 4.11: Change in Math Proficiency: All Schools

The only consistent trend across grade levels involves the percent of advanced students. In all grade bands and all subjects, the control schools exhibit more favorable changes in the percentages of students categorized as advanced. Within grade bands, the high school data exhibits troubling trends when compared to the control schools in all student categories (% P/A, % Below Basic, and % Advanced). The Reading/Language Arts results for Elementary and Middle school students show encouraging trends in the percentage of students who are testing as proficient or advanced as well as decreasing the relative number of students testing as below basic.

The inconsistencies in the proficiency trend data makes it difficult to come to any conclusions. In some cases, the proficiency results contradict the information communicated in the growth data. For example, high school math results appear very strong in the growth data, but exhibit the opposite trends as expected in the proficiency data. It is possible that the proficiency data is too cohort-dependent to provide meaningful data. These trends will continue to be monitored in the future in order to better determine their value in the overall analysis of the STC.

Conclusions and Considerations

The largest take-away from this analysis is that the Knox County School Technology Challenge (SCT) is very much a work in progress. The data from the Technology Integration Matrix and the Personalized Learning Environment rubric suggest that the level of implementation in year one was not yet sufficient to fundamentally change the classroom experience in STC schools.

Certainly, there are encouraging trends in the data. The growth data, at least directionally, illustrates some of the promise of what could be achieved with deeper implementation. Some of the trends that show less promising trends could be attributed to dueling dynamics. For example, purposeful comparison of teacher level TVAAS results and observation scores generally excludes novice teachers. These teachers may be experiencing gains in these metrics that are driven more by their years of experience than by the deployment of technology. However, the teacher interviews indicated that the novice teachers were some of the most enthusiastic adopters of the STC. Different analysis methods may be required to address this issue during future studies.

The unfavorable change in some of the metrics in STC schools should also be considered in context. The implementation of the STC initiative can represent a monumental shift in a teacher's approach to their craft. Many teachers noted the difficulty of dealing with the magnitude of the change during focus group interviews. It is possible that unfavorable results are the classic "implementation dips" that oftentimes accompanies a change of this magnitude. It is important, however, to not simply write-off all unfavorable outcomes as resulting from implementation dip. For example, this analysis found a statistically significant

negative relationship between TVAAS index and the "Authentic" indicator on the Technology Integration Matrix. This relationship may be evidence of the implementation dip, but it is important to also consider other explanations. It is possible that teachers were focused on real-world problems and projects that were interesting to the students but not necessarily aligned with the curriculum. It is also possible that these projects were either aligned with too few standards or took too much time to implement that instructional time was lost for covering other standards. Field research will need to be conducted to answer some of these lingering questions.

There is a lot of opportunity to help tie the STC to other Knox county strategic goals. Knox County strives to be a learning, collaborative environment. However, STC teachers either did not see the value or perceive that they did not have the opportunity for teacher-to-teacher collaboration between STC schools. KCS should take a stronger role in helping to build a community of collaboration between the teachers at the STC schools.

The other strategic opportunity involves the KCS magnet program. Many elementary teachers were lamenting the fact that some of their 5th grade students would be transitioning to a Middle school that was not a participant in the STC. This is an opportunity for KCS to market Vine Middle School's technology magnet program to strengthen Vine's magnet program. Similarly, 1:1 L&N STEM Academy can serve as the magnet for the students attending an STC middle school but zoned for a non-STC high school. Feeder patterns should also be considered if there is an eventual expansion of the SCT initiative, since Vine Middle and L&N STEM may not be able to absorb all the demand for a 1:1 environment. The true demand for the STC concept may be explored in the future by including student transfer rates (both in and out) of STC schools in the next formative evaluation cycle.

Going forward, principals have been encouraged to collect TIM and PLE rubric scores from their teachers on a regular basis. Feedback from the rubric will allow the staff at STC schools to gain insight into the breadth and depth of their implementation. Also, it is one of the few leading indicators available for the STC initiative.

The district has already adjusted professional development offerings based on the feedback from the focus groups. Technology coaches are receiving targeted training that focuses on each indicator of the PLE and TIM rubrics. These coaches have been tasked with providing building-level training based on these district-led professional development offerings.

Modeling work will begin to provide another mid-year feedback loop in the elementary schools with the STAR Renaissance tests. These formative assessments can be used to gauge student performance in STC schools against students in the control schools.

Meetings will also occur with district leadership to define which indicators we consider key to the success of the program. The goal of this activity would be to maximize the signal to

noise ratio to ensure that future decisions are being made based upon the most important and reliable information.

The final next step is to get back in front of the STC teachers for follow-up focus groups. The current plan is to present this information to the STC teachers for comment and to help answer some of the lingering questions. Discussions will also focus around how implementation has changed in year two of the STC and the effectiveness of the revamp professional development offerings.

Individual Learning Cycles

As part of the Effective Educators initiative, instructional coaches work one-on-one with certain teachers via individual learning cycles (ILCs). ILCs are part of the overall learning cycle for teachers that are meant to focus on Common Core State Standards transition, literacy instruction, and TEAM instructional indicators by providing teachers high-quality learning experiences and intensive classroom support. Learning cycles embody research-based best practices for professional development: support over time, on-going follow up, targeted and specific support, and a gradual release of responsibility from the coach to the teacher.

ILCs are opportunities for the coach and teacher to craft a personalized plan to help improve the teacher's instructional delivery and student outcomes. The coach and teacher collaboratively determine the focus of the ILC and coaching plan. ILCs typically include: modeling, co-teaching, co-planning, observing, and providing feedback. Among the key outcomes of the learning cycle is for teachers to learn instructional strategies that will impact student performance. Since those outcomes are not strictly defined, we analyzed teacher evaluation data and student growth data as the units of analysis for this evaluation, in addition to qualitative perception data.

Methodology: Data Used

How are teachers selected to participate in an ILC? Generally, it is teachers who earned a Level of Effectiveness or individual growth score of Level 1 or 2. In practice, principals generate a list of teachers in each building that qualify for ILCs based on their classroom observation scores, overall effectiveness scores, or if they are new teachers. Additionally, some teachers self-select to participate in ILCs. The breakdown of how teachers were selected for ILCs in school year 2013-2014 (SY1314) is shown in Table 5.1 based on data provided by the instructional coaches leading ILCs.

Table 5.1: ILC Referral Distribution

Basis for ILC Referral	Count	% of Total
New Teachers	11	2.4%
Principal/Scores	437	94.4%
Self-selected	15	3.2%

This is not a clean-cut view of how teachers are selected for ILCs because there is overlap in referrals. For example, a new teacher may have qualified for an ILC through scores, but was only recorded under the "Scores" category. Using this list provided by instructional coaches, we were able to gather data for a majority of the ILC participants.

District-wide, 42% of teachers who earned a Level of Effectiveness score of 1 or 2 participated in an ILC in SY1314. Table 5.2 shows the percent of ILC Teachers by Level of Effectiveness from the district-wide pool.

Table 5.2: District-wide Level of Effectiveness Score Distribution

	SY1213 Level of Effectiveness					
	1	2	3	4	5	Total
Active KCS Teachers	19	316	816	1600	895	3646
ILC Teachers	16	125	119	61	6	327
ILC % of Column Total	**84.2%**	**39.6%**	**14.6%**	**3.8%**	**0.7%**	**9.0%**

In sum, there were only three Level 1 teachers who did not participate in an ILC. Conversely, almost 60% of Level 2 teachers did not work directly with an instructional coach via ILC.

Level of Effectiveness scores are based, in part, on classroom observation data and individual growth scores. The Level of Effectiveness, as well as individual growth scores, are some of the components considered when placing teachers in ILCs. Therefore, we analyzed two data points: classroom observation scores and TVAAS gains (from which the individual growth score is derived).

- Classroom Observation Scores
  - These scores are from classroom observations and professionalism ratings based on role-specific rubrics (general educator, library media specialist, and school services personnel).
- Individual Growth Scores and TVAAS Gains
  - These scores are the teachers' value-added scores (TVAAS) based on three years of data.

Using teacher license numbers, we were able to match evaluation data with approximately 93% (430 of 463) of the teachers enrolled in ILCs. Table 5.3 summarizes the statistical tests conducted in this report based on the data point and type of teacher. We also included perception data collected from surveys.

Table 5.3: Statistical Tests Used in ILC Analysis

Data Point	Teacher Type	Paired t-test	Independent Samples t-test	Chi-square
Observation Scores	ILC-only	✓		
	ILC and Control		✓	✓
	ILC and Non-ILC Apprentice		✓	
Individual Growth Scores & TVAAS Gains	ILC-only	✓		
	ILC and Control		✓	✓
	ILC and Non-ILC Apprentice		✓	

A note on controlling for years of service

Teacher growth data is dependent upon years of experience in the first five years of teaching—new teachers have high rates of change and growth. In order to control for the effect of years of service on performance, we included an equal amount of new teachers in our control groups to minimize this effect as much as possible. However, we were not able to find an adequate number of two and three-year teachers who earned an individual growth score of Level 1 for use in that control group. Therefore, the analysis was limited to those teachers in the ILC and control group who earned a score of 2 or greater. This instance of limited analysis is noted clearly in the methodology and results sections below.

Methodology: ILC Teachers Compared to Themselves

We were able to collect two years of data (SY1213 and SY1314) for 317 ILC teachers. We used paired t-tests on classroom observation scores, individual growth scores, and TVAAS gains. We also conducted a sub-analysis of only Level 1 and 2 ILC teachers by way of paired t-tests. This sub-analysis was done to shed light on the intended target of ILCs.

Methodology: ILC Teachers and Control Groups

While comparing the ILC teachers to themselves is useful, it is also important to know if participation in the ILC helped differentiate them from similarly situated peers. Therefore, teachers who participated in an ILC and had SY1213 and SY1314 evaluation data were matched with a control group of teachers who were not in an ILC but had similar years of service and similar SY1213 evaluation data. We created two control groups in order to align the ILC (treatment) group as closely as possible to the control groups based on classroom observation scores and individual growth scores.

Methodology: ILC & Control Classroom Observation Scores

Figure 5.1 shows the distribution of SY1213 classroom observation scores for the ILC treatment and control groups.

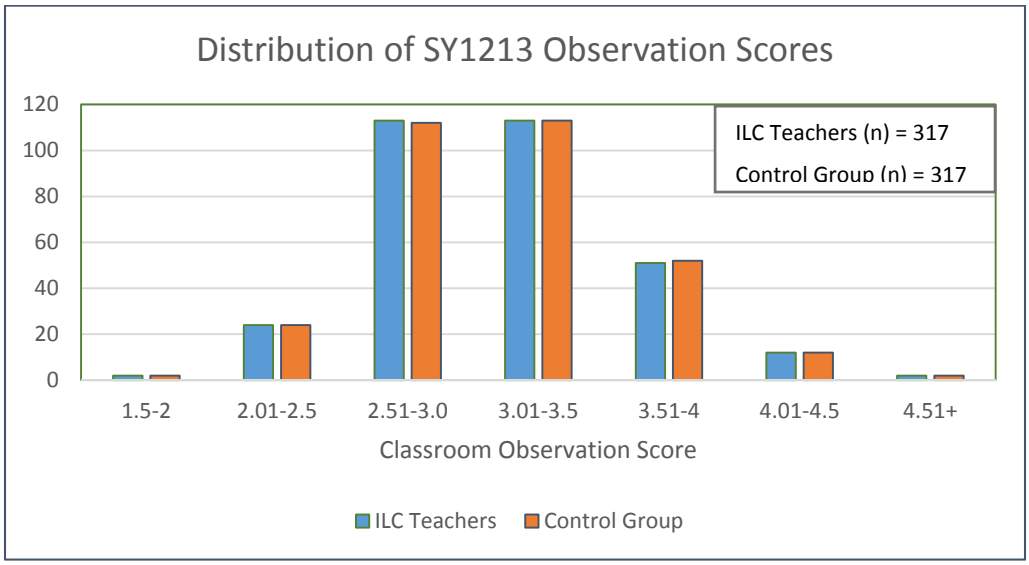


Figure 5.1: Distribution of SY1213 Classroom Observation Scores in Treatment and Control Groups

An analysis of observation scores is difficult due to the school-to-school variation in the mean classroom observation score. To remove the school-to-school variation in the mean observation score, the difference between teachers' classroom observation scores and their schools' mean observation score was calculated. The following is an algebraic expression of how this difference (delta Δ) was calculated.

$$\Delta_{Obsv\ Score} = (Teacher's\ Obsv\ Score - Mean\ School\ Obsv\ Score)_{SY1314} - (Teacher's\ Obsv\ Score - Mean\ School\ Obsv\ Score)_{SY1213}$$

A t-test was done on the mean change (Δ) of teachers in ILCs and teachers not in an ILC. The null hypothesis tested was that the mean change in classroom observation scores from SY1213 to SY1314 was no different for both ILC teachers and the control group. A chi-squared test was used to evaluate if the distribution of scores that improved or decreased was equal among ILC teachers and the control group.

Apprentice Only

A sub-analysis was also conducted of apprentice teachers only, since ILCs are intended to instill instructional excellence in new teachers. We compared the distance to the mean classroom observation score for ILC apprentice teachers to non-ILC apprentice teachers using an independent samples t-test.

Methodology: ILC & Control Group Individual Growth Scores

Figure 5.2 shows the distribution of SY1213 individual growth scores for the ILC treatment and control groups.

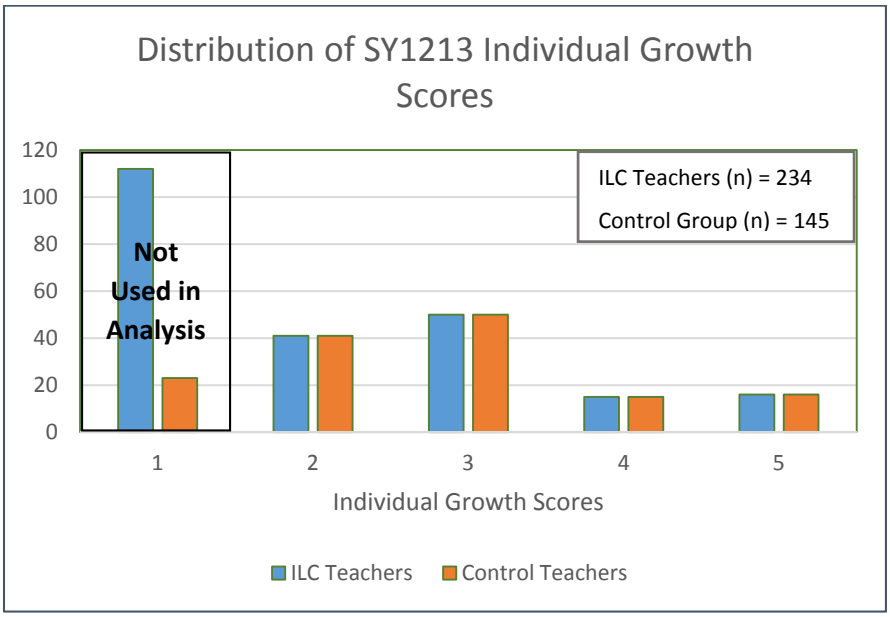


Figure 5.2: Distribution of SY1213 Individual Growth Score by Level in ILC and Control Groups

In order to appropriately control for scores of Level 1 and apprentice teachers, we had to limit the evaluation of ILC teachers and the control group teachers to those with individual growth scores of Levels 2-5. The control group was created based on the level growth score (2-5), but the analysis was conducted using TVAAS gains. A t-test was conducted to test whether the mean TVAAS gain from SY1213 to SY1314 between the ILC teachers and the control group was no different. A chi-squared test was used to determine if the distribution of TVAAS gains that improved or decreased was equal among ILC teachers and the control group.

Apprentice Only

A sub-analysis was also conducted of apprentice teachers only, since ILCs are intended to instill instructional excellence in new teachers. We compared mean TVAAS gains of ILC apprentice teachers to non-ILC apprentice teachers using an independent samples t-test.

Results: ILC-Only Classroom Observation Scores

We conducted a paired t-test on the distance to the mean classroom observation score for ILC teachers in SY1213 and SY1314. The results are in Table 5.4.

Table 5.4: Paired t-test Results of ILC-Only Analysis of Classroom Observation Scores

	SY1213	SY1314
Mean Distance to Average Observation Score	-0.35	-0.33
Observations	317	317
p-value	0.470	

Since the p-value is greater than 0.05, we cannot conclude there is a statistical difference between the means. However, the negative means indicate that the ILC teachers earned classroom observation scores that were below the average classroom observation scores in their schools, though the mean distance to the average score shrunk by 0.02 points after treatment.

Results: Observations Scores of Intended ILC Target Teachers

We also conducted a paired t-test on only those ILC teachers who earned an observation score of 3.0 or less (the intended target of ILCs since the focus of ILCs is supposed to be teachers with low observation and effectiveness scores). The results are in Table 5.5.

Table 5.5: Paired t-test Results of ILC-Only Teachers with SY1213 Observation Scores of 3.0 or Less

	SY1213	SY1314
Mean Distance to Average Observation Score	-0.68	-0.50
Observations	139	139
p-value	**0.000***	

The p-value is 0.0, which indicates there is a statistically significant difference between the mean distance to the average observation score from one year to the next. A visual inspection of the means showed that although the mean classroom observation score for this subset of ILC teachers was below the average score in their buildings, the gap between their mean and the school mean decreased in a significant way (0.18 points) after treatment. When we further narrowed our analysis to those critical target ILC teachers who had a starting observation score of less than 2.55, the gap between their mean score and the school mean score decreased 0.32 after treatment, and the difference was statistically significant. See Figure 5.3.

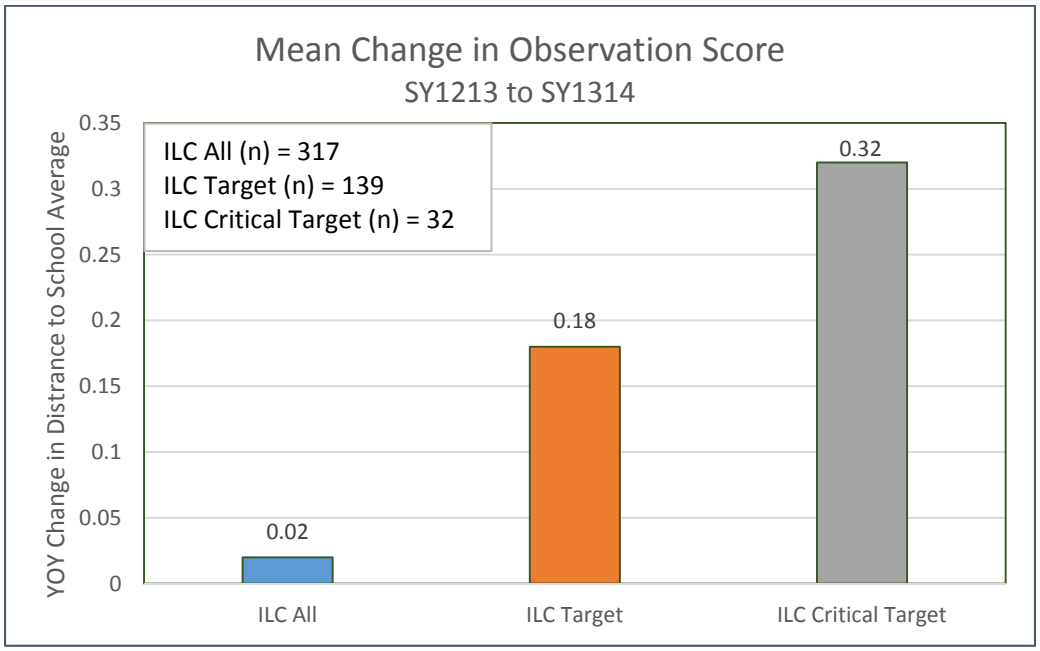


Figure 5.3: Mean Change in Distance to School Observation Score Average, 2012-2014

Results: ILC-Only Individual Growth Scores

There were 234 ILC teachers who generated an individual growth score in SY1213 and SY1314. Half of them did not change levels. Approximately 27% of them improved their individual growth score, while 23% did worse after the ILC (see Figure 5.4).

SY1314 Individual Growth Score

SY1213 Individual Growth Score: Level	1	2	3	4	5
1	70	18	17		7
2	17	13	8	3	
3	11	5	26	6	2
4	2	1	7	3	2
5	1	1	2	6	6

Same	50%
Better	27%
Worse	23%

Figure 5.4: Distribution of Individual Growth Scores by Year

The results of the paired t-test are in Table 5.6.

Table 5.6: Paired t-test Results of ILC-Only Analysis of TVAAS Gains

	SY1213	SY1314
Mean TVAAS Gains	-3.5	-2.1
Observations	234	234
p-value	**0.000***	

Since the p-value is 0.00, we can conclude there was a statistical difference between the mean TVAAS gain in SY1213 and SY1314. Visual inspection of the means indicated that ILC teachers still had a mean negative gain, but the mean did increase from SY1213 to SY1314.

Results: Individual Growth Scores of Intended ILC Target

Since ILCs are geared toward teachers who earned individual growth scores of 1 and 2 in SY1213, we calculated rates of improvement for only that intended target population. After treatment, over a third (35%) of those ILC teachers earned better individual growth scores, while only 11% earned lower individual growth scores. These results are bolstered by a sub-analysis of TVAAS Gains of ILC teachers who fell within the intended ILC range, see Table 5.7.

Table 5.7: Paired t-test Results of ILC-Only Teachers with SY1213 Individual Growth Scores of Level 1 or 2

	SY1213	SY1314
Mean TVAAS Gains	-6.0	-2.6
Observations	153	153
p-value	**0.000***	

Because the p-value is 0.00, we can conclude there is a statistically significant difference between the mean TVAAS gain in SY123 and SY1314. Visual inspection of the means indicated that although this subset of ILC teachers still had a mean gain below the growth standard, they improved their mean TVAAS gain by 3.4 points. This improvement is further emphasized when considering just those critical target ILC teachers who earned an individual growth score of Level 1 in SY1213, see Table 5.8.

Table 5.8: Paired t-test Results of ILC-Only Teachers with SY1213 Individual Growth Scores of Level 1

	SY1213	SY1314
Mean TVAAS Gains	-6.8	-2.5
Observations	112	112
p-value	**0.000***	

Since the p-value is 0.00, we can conclude there is a statistically significant difference between the mean TVAAS gain in SY1213 and SY1314 for those teachers who earned an

individual growth score of Level 1 in SY1213. Visual inspection of the means indicated that although this subset of ILC teachers had a mean gain below the growth standard, they improved their mean gain by 4.3 points. That is almost one point greater than the target ILC teacher population and three points more than the full ILC teacher population, see Figure 5.5.

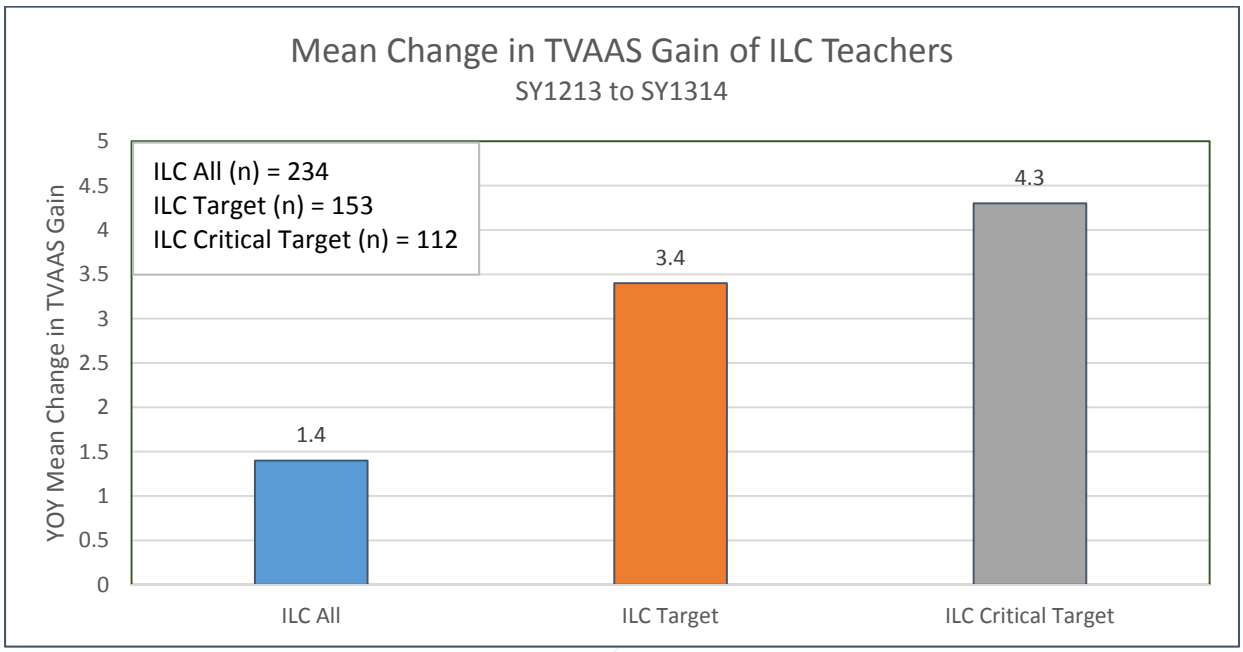


Figure 5.5: Mean Change in TVAAS Gain of ILC Teachers, 2012-2014

Results: ILC & Control Classroom Observation Scores

In order to control for the variation in observation scores across the schools, we tested the hypothesis that the delta (the mean year over year change in the distance to the average classroom observation score) was no different between ILC teachers and the control group. Table 5.9 shows the results of the independent samples t-test.

Table 5.9: t-test Results of ILC and Control Group Analysis of Classroom Observation Scores

	ILC Teachers	Control Group
Mean Change in Delta Δ	0.017	0.169
Observations	317	317
p-value	**0.000***	

Since the p-value is 0.00, we can conclude that there is a statistically significant difference between the mean delta of both groups (ILC teachers and control group). These results are augmented by the outcome of the chi-square test used to evaluate if the distribution between the two groups was equal. More teachers in the control group earned higher classroom observation scores than the ILC teachers.

Apprentice Only

A similar evaluation was conducted for apprentice teachers only, comparing ILC to non-ILC apprentice teachers.

Table 5.10: t-test Results of Apprentice Teacher Analysis of Classroom Observation Scores

	ILC Apprentice Teachers	Non-ILC Apprentice Group
Mean Distance to Average Classroom Observation Score	0.19	0.20
Observations	98	365
p-value	0.701	

Since the p-value is greater than 0.05, we cannot conclude that there is a statistically significant difference between the two groups. Directionally, however, the mean difference of both groups was roughly 0.2 points over the average observation score in their school.

Results: ILC & Control Group Individual Growth Scores

We tested the hypothesis that the mean TVAAS gain was no different between ILC teachers and the control group. Since we were not able to control for years of service for those teachers who earned an individual growth score of Level 1, we limited this analysis to those with a score of 2 or greater. Table 5.11 shows the results of the independent samples t-test.

Table 5.11: t-test Results of ILC and Control Group Analysis of TVAAS Gains

	ILC Teachers	Control Group
Mean TVAAS Gain for those with score of 2+	-1.46	-0.15
Observations	121	130
p-value	**0.050***	

Since the p-value is 0.05, we can conclude with 95% confidence that the mean TVAAS gain among the two groups is different. Based on visual inspection of the means, the mean TVAAS gain of ILC teachers was worse than the control group teachers. The control group teachers showed negative growth, though not as profound as ILC teachers.

Table 5.12 shows the distribution of change in each group's mean TVAAS gains. (Please note this is limited to teachers who earned a Level 2 or above since we were unable to control for apprentice teachers who earned a Level 1.)

Table 5.12: Distribution of Change for ILC and Control Group Analysis of TVAAS Gains

Group	Level Improved		Level Decreased	
ILC Teachers	46	38%	75	62%
Control Group	59	48%	63	52%

The percent of ILC teachers whose TVAAS gain decreased was more than half the total. The chi-square test resulted in a p-value of less than 0.05, so we can conclude that there is a statistically significant difference in the distribution between the two groups. In this case, the control group increased their scores at a higher rate than the ILC teachers.

Apprentice Only

We conducted a similar evaluation of mean TVAAS gain of apprentice teachers only, comparing ILC apprentice teachers to non-ILC apprentice teachers. Results of the independent samples t-test are below.

Table 5.13: t-test Results of Apprentice Teacher Analysis of TVAAS Gains

	ILC Apprentice	Non-ILC Apprentice
Mean TVAAS Gain	1.76	1.75
Observations	74	37
p-value	**0.991**	

Since the p-value is greater than 0.05, we cannot conclude that there is a statistically significant difference between the two groups. Directionally, however, both types of apprentice teachers showed positive mean gains.

Perception Data

TELL Survey

The TELL (Teaching, Empowering, Leading and Learning) Tennessee Survey was launched in 2011 and is administered every two years. According to its website, the survey provides educators with data to facilitate school improvement. It includes questions from a range of topics, including teacher leadership, use of time, professional development, and instructional practices and support. The results from the TELL survey are especially helpful because the first administration was prior to the revamp of the coaching model in the district. In 2011, almost 73% of teachers responded to the survey, while in 2013 only 44% responded. The questions were not specific to ILCs, but rather about instructional coaches, time for collaboration, and access to various resources. Inferences about ILCs were made based on the responses.

In sum, teachers surveyed in 2013 reported that: they had less time to collaborate but more access to professional support personnel; professional development did not enhance their ability to help improve student learning and it did not deepen teacher content knowledge; they need less professional development; and fewer new teachers reported receiving support in 2013 than in 2011. Table 5.14 shows the growth in certain areas from the first to the second administration of the survey. The results are based on favorable responses (agree or strongly agree).

Table 5.14: TELL TN Survey Results, 2011 and 2013

Question	2011	2013	Growth	Implications
Teachers have time available to collaborate with colleagues.	70%	60%	decrease	The decrease here indicates that teachers reported having less time for the specified tasks.
The non-instructional time provided for teachers in my school is sufficient.	53%	48%	decrease	
Collaborative planning time (time spent: 0-3 hours/week)	76%	84%	increase	The increase here indicates that more teachers reported spending between 0-3 hours per week on the specified tasks.
Professional development (time spent: 0-3 hours/week)	81%	83%	increase	
Teachers have sufficient access to appropriate instructional materials.	81%	85%	increase	The increase here indicates that more teachers reported having sufficient access to the specified tasks.
Teachers have sufficient access to a broad range of professional support personnel.	80%	82%	increase	
Teachers receive feedback that can help them improve teaching.	83%	84%	increase	The increase in here indicates that teachers reported that school leadership adequately executed or supported the specified tasks.
The school leadership makes a sustained effort to address teacher concerns about: professional development.	76%	78%	increase	
The school leadership makes a sustained effort to address teacher concerns about instructional practices and support.	85%	87%	increase	
Sufficient resources are available for professional development in my school.	81%	85%	increase	The mixed results here indicate that while certain aspects of professional development increased (resources, time spent, differentiation), there were also areas that teachers rated more favorably in 2011 (PD increasing content knowledge and improving student knowledge).
An appropriate amount of time is provided for professional development.	78%	80%	increase	
Professional development is differentiated to meet the needs of individual teachers.	59%	62%	increase	
Professional development deepens teachers' content knowledge.	71%	70%	decrease	
Teachers are encouraged to reflect on their own practice.	90%	94%	increase	
Follow up is provided from professional development in this school.	64%	68%	increase	
Professional development provides ongoing opportunities for teachers to work with colleagues to refine teaching practices.	75%	78%	increase	
Professional development is evaluated and results are communicated to teachers.	56%	54%	decrease	
Professional development enhances teachers' ability to implement instructional strategies that meet diverse student learning needs.	80%	81%	increase	
Professional development enhances teachers' abilities to help improve student learning.	84%	75%	decrease	
In which of the following areas do you need professional development to teach your students more effectively: content area	37%	34%	decrease	The decrease here indicates that teachers reported needing less professional development in the specified areas.
In which of the following areas do you need professional development to teach your students more effectively: differentiating instruction	63%	61%	decrease	
In which of the following areas do you need professional development to teach your students more effectively: methods of teaching	42%	37%	decrease	
In which of the following areas do you need professional development to teach your students more effectively: classroom management techniques	35%	31%	decrease	
Teachers work in professional learning communities to develop and align instructional practices.	89%	93%	increase	The mixed results here indicate that while teachers felt PLCs helped, not all supports translated to improved instructional practices.
Provided supports (i.e., instructional coaching, PLCs, etc.) translate to improvements in instructional practices by teachers.	83%	82%	decrease	
As a beginning teacher, I have received the following kinds of support: a formally assigned mentor	87%	82%	decrease	The decrease here indicates that new teachers reported receiving less support in these areas than reported in 2011.
As a beginning teacher, I have received the following kinds of support: common planning time with other teachers	80%	72%	decrease	
As a beginning teacher, I have received the following kinds of support: access to PLCs where I could discuss concerns with other teachers	84%	82%	decrease	

District Survey

The Department of Research, Evaluation, and Assessment (REA) administered a district-wide survey, in part, to gather perception data about instructional coaches. The survey was sent in May 2014 to a random selection of school administrators, classroom teachers, instructional assistants, instructional coaches, and other certified personnel. Because this survey was first administered in 2014, we are unable to compare previous results or show any growth.

To capture teacher responses, we calculated favorable responses by combining "agree" or "strongly agree" responses. The questions were not specific to ILCs, but rather about instructional coaches, topics covered, and changes to instructional practices as a result of instructional coaches. Inferences about ILCs were made based on the responses. Approximately 342 teachers answered the questions related to instructional coaches: roughly 40% were elementary school teachers, 21% were middle school teachers and slightly more than 31% were high school teachers (the remaining 8% were teachers of multiple grade levels, Pre-K, or unspecified grades).

Based on the responses, respondents indicated they felt the most impact from instructional coaches in the areas of supporting collaboration (58% favorable responses) and providing access to and encourage use of different and new resources (53% favorable responses). There were less favorable responses about the impact instructional coaches had on learning the evaluation rubric or changing knowledge of instructional practices (both had 36% favorable responses). Overall, elementary respondents had a more positive perception of the impact instructional coaches have, whereas secondary respondents tended to respond less favorably. A summary of the favorable responses can be found in Table 5.15

Table 5.15: KCS District Survey Results, 2014

Question	% of Favorable Responses by Teacher Level			
	Overall	Elementary	Middle	High
The instructional coaches led teachers to think about an aspect of their teaching in a new way**.**	53%	63%	58%	35%
The instructional coaches encouraged teachers to pay closer attention **to particular things that were being taught.**	52%	65%	51%	32%
The instructional coaches led teachers to seek out additional information or other resources**.**	53%	66%	52%	36%
The instructional coaches encouraged collegiality and collaboration **among teachers.**	58%	72%	52%	43%
The instructional coaches led teachers to question their beliefs and assumptions about which teaching methods **work best with students.**	42%	50%	45%	28%
As a result of interactions with instructional coaches, have there been changes in classroom management practices**?**	21%	22%	14%	21%
The instructional coaches led teachers to modify or improve the lesson planning process**.**	50%	60%	49%	34%
As a result of interactions with instructional coaches, have there been changes in teacher knowledge and understanding of instructional practices**?**	36%	46%	22%	29%
As a result of interactions with instructional coaches, have there been changes in teacher knowledge and understanding of the evaluation rubric**?**	36%	43%	25%	33%
As a result of interactions with instructional coaches, have there been changes in the use of student data in instructional planning**?**	40%	52%	33%	26%
As a result of interactions with instructional coaches, have there been changes in the use of differentiated instructional strategies**?**	40%	49%	32%	30%

Coaching Focus Groups

In the spring semester of SY1314, the REA hosted a series of eight focus groups with every type of instructional coach (secondary math, elementary, literacy, etc.). They were asked six questions related to the coaching model, their needs, and their perceived impact. These questions also led to other questions and open dialogue. Several themes emerged, but the

one that was repeated in every group was: principals make all the difference in a coach's effectiveness in a school. Some other feedback received that is relevant to ILCs is below.

- Several coaches reported that principals did not always share the areas of refinement from a teacher's evaluation, which would help inform the ILC plan. While evaluation data is confidential and sensitive, providing coaches with some appropriate information may help focus the ILC on those areas of refinement.
- Many coaches reported that some teachers were led to believe ILCs are a punitive measure or were not given notification they would be enrolled in an ILC. Coaches reported this lack of information on the front-end caused mistrust and discontent on the part of the teacher in the initials weeks of the first cycle.
- Several coaches reported working in an ILC with a teacher from a different content area (literacy coach working with a special education teacher, for example), which may have an adverse effect.
- Coaches requested additional training in adult learning for ILCs, as well as specialized content areas.
- Some coaches reported ILCs are more reflective of a coach's impact than other coaching tasks (PLCs, professional development, assessment-development, etc.) since it is tied to one teacher and the teacher data is easy to track. Some coaches also recommended evaluating growth in observation scores.

The coaching model outcomes call for coaches to receive intensive training in content areas, professional delivery models, TEAM, and cognitive coaching competencies. The feedback from coaches suggests that these outcomes have not been fully reached.

Conclusions and Considerations

When considering the results of the ILC analysis, it is important to underscore a few issues:

- Observation requirements changed mid-year in SY1314. In Fall 2013, all professionally licensed teachers had unannounced observations. However, based on a change in policy by the Superintendent, some teachers opted to have announced observations in Spring 2014. The initial practice of unannounced observations may have contributed to a change in scores.
- There may be other confounding issues that contribute to the results, such as the changes to and implementation of the TEAM evaluation system and the use and structure of professional learning communities (PLCs) in schools. It is difficult to control for or distill their effect on the ILC evaluation.

The results of this analysis suggest that ILCs have positive impacts for those teachers who qualify for ILCs based on low scores. Based on the quantitative metrics we used, the strongest indicator of improvement was found among those ILC teachers who qualified for ILCs. There was no evidence to suggest that apprentice teachers who were enrolled in an

ILC performed any better than their non-ILC apprentice peers. Table 5.16 summarizes the overall results of the ILC evaluation.

Table 5.16: Overall Results of SY1314 ILC Evaluation

Unit of Analysis	Classroom Observation Scores	TVAAS Gains
Intended ILC Target	Increased*	Increased* but still below growth standard
All ILC Teachers	Below school average	Increased* but still below growth standard
ILC & Control Group	Control group outperformed ILC*	Control group outperformed ILC*
ILC & Non-ILC Apprentice	Increased but no discernible difference from Non-ILC Apprentice Teachers	Increased but no discernible difference from Non-ILC Apprentice Teachers

*denotes statistical significance

Overall, the mean observation score of ILC teachers did not improve but ILC teachers did increase their mean TVAAS gain—though it was still below zero (i.e., below the growth standard). When compared to the control group, ILC teachers did not perform as well as their similarly situated peers. When looking just at apprentice teachers, ILC and non-ILC apprentice teachers' mean outcome data was almost identical. However, when looking *only* at those teachers who are the *intended target* of an ILC, that is—they earned Level of Effectiveness or individual growth scores of Level 1 or 2, there was clear improvement in their mean evaluation data.

If the district chooses to continue monitoring and investing in instructional coaches at the central level, it may wish to narrow the focus and approach of how instructional coaches reach certain teachers via ILC, as well as modify certain data collection practices. Changes to consider are as follows:

- The district should consider limiting ILCs to those teachers with a Level 1 or 2 score, since that group of ILC teachers showed significant improvement.
- Only 40% of teachers who earned a Level of Effectiveness score of 2 were enrolled in an ILC. Given that Level 1 and 2 teachers had statistically significant improvement in their mean data, it would behoove the district to encourage Level 2 enrollment in ILCs.
- Conversely, the district should consider limiting the number of effective teachers in ILCs. Almost 20% of teachers who were selected to participate (not self-selected) in an ILC in SY1314 were professionally licensed, considered "effective" (they had a Level of Effectiveness score of 3, 4 or 5), and had classroom

observations score of 3.0 or higher in SY1213. It is unclear why these effective teachers were enrolled in an ILC—even if they self-selected to enroll—they should have been directed to different resources rather than an ILC, such as PLC leadership or the mentoring program. ILCs did not have the intended effect on this group of teachers.

- The district may wish to bolster its current mentoring program (funded by the Great Schools Partnership) by connecting new teachers with highly effective teachers in their buildings. This would free up coaches' time for professionally licensed teachers who have earned scores below expectations.
- The district should continue to build upon its current (SY1415) practice of providing instructional coaches with additional and consistent TEAM rubric, adult learning, and content-area training opportunities.
- The district may wish to develop best practices surrounding the ILC process, including (but not limited to)
    - A meeting between the principal and coach to discuss the areas of refinement (from the evaluation) for each teacher in an ILC so that the coach can plan to address those areas.
    - An initial meeting between the principal, teacher, and coach should be scheduled prior to or during the first meeting of the ILC to establish a positive and productive perception about the process. This would also provide an opportunity for the teacher to ask questions about the process if s/he is unfamiliar with it.
    - Coaches should, whenever possible, work with teachers within their content specialty.
- Since the referral process is based on multiple data points and several decision-makers, the data collection form should be completed such that there is clear delineation why a teacher is selected for an ILC.
- In addition to ensuring accurate referral data is collected on the ILC enrollment form, data on how often teachers and coaches are meeting should be uniform and collected centrally.

Any iteration of this evaluation should include focus groups and/or individual surveys of teachers who participated in an ILC in order capture their perception of the process and its utility. Any additional metrics that accurately capture the key outcomes of the coaching model and ILCs may also need to be revisited with the REA and the Curriculum and Instruction department.

TAP

In SY0607, Knox County Schools implemented TAP: The System for Teacher and Student Advancement (previously known as the Teacher Advancement Program) in four of its highest needs schools. The program was expanded in SY1112 and now includes the eighteen schools listed below.

Table 6.1: TAP Schools

Type	School	Level
Original TAP School	Lonsdale Elementary	Elementary
Original TAP School	Pond Gap Elementary	Elementary
Original TAP School	Holston Middle	Middle
Original TAP School	Northwest Middle	Middle
TAP School	Belle Morris Elementary	Elementary
TAP School	Dogwood Elementary	Elementary
TAP School	East Knox County Elementary	Elementary
TAP School	Ritta Elementary	Elementary
TAP School	Sarah Moore Greene Elementary	Elementary
TAP School	Spring Hill Elementary	Elementary
TAP School	West Haven Elementary	Elementary
TAP School	West View Elementary	Elementary
TAP School	Austin-East High	High
TAP School	Carter High	High
TAP School	South-Doyle High	High
TAP School	Carter Middle	Middle
TAP School	South-Doyle Middle	Middle
TAP School	Vine Middle	Middle

TAP is a comprehensive school reform model developed by Lowell Milken. The primary goal of the TAP program is to increase teacher recruitment, retention, motivation, practices, and performance. The Milken foundation identifies four elements to the successful implementation of the TAP initiative.

- Multiple career paths – Master and mentor teachers are chosen based on their strong classroom performance and their expert-level curricular knowledge. Master and mentor teachers are part of the school's leadership team and are instrumental in goal setting and monitoring, as well as providing feedback and support to other teachers. They create a network of support comprised of career educators.
- Ongoing professional growth – The traditional school day is modified so that teachers can collaborate in planning, professional learning, and mentoring in clusters. The

focus of this collaboration is to provide student-centered support via master and mentor teachers. The level of collaboration is intended to drive student academic achievement.

- Instructional-based accountability – Teachers in a TAP school are evaluated (by multiple observers) in their teaching skills as well as their knowledge and responsibilities around TAP performance standards. The heart of the evaluation process is in the open discussion regarding teacher strengths and weaknesses and is the driving force in strengthening teacher effectiveness.
- Performance-based compensation – Teacher salaries are augmented by additional compensation. This allows more effective teachers to be compensated above and beyond the traditional district pay structure. Teachers are also compensated for increasing their job responsibilities as master or mentor teachers in a TAP school. Performance-based compensation is tied directly to student growth to help reward high-quality practice.

Early KCS studies of the TAP initiative suggested that student performance on end-of-year assessments may increase after TAP implementation. There were mixed findings on the sustainability of these results as schools moved further from the date of implementation.

However, the landscape of the district has changed considerably since those early studies. In SY1112, non-TAP schools were mandated by the state of Tennessee to conduct annual evaluations of their certified staff through the TEAM process (Tennessee Educator Acceleration Model). The TEAM rubric is virtually identical to the TAP observation rubric.

TEAM schools have also since been staffed with TEAM lead teachers. These lead teachers conduct some of the required teacher evaluations and provide peer-to-peer feedback to drive instructional improvement. It should be noted that the responsibilities of TEAM lead teachers differ from TAP master and mentor teachers. TEAM lead teachers function more wholly within the confines of classroom observations. There are more master and mentor teachers per TAP school available to provide deeper professional development.

Finally, KCS opted to institute a bonus compensation program based on observation scores and student growth results in TEAM schools. The district has been committed to creating collaborative professional learning communities (PLCs) in schools to function within the same theory of action as TAP's cluster strategy. These changes have helped to close (but not eliminate) the operational gaps between TAP and non-TAP schools.

This evaluation is focused on comparing performance in TAP schools to performance in similar TEAM schools.

Methodology

The difficulty in the analysis involves creating a suitable control group. High-needs schools serving traditionally underperforming students were targeted for the adoption of TAP. It is difficult to find schools that are ideal comparisons to the TAP schools based on demographics and past academic performance. The difficulty is most acute for middle schools.

Control schools were chosen from hierarchal clustering data. Control schools are the nearest TEAM neighbors to each TAP school (when considering key demographics). The four key demographic groups are defined by the Tennessee Department of Education Accountability system (percent of: free and reduced price lunch students; Black, Hispanic or Native American students; special education students; and English Language Learners). The pool of control schools is listed in Table 6.2 below.

Table 6.2: Comparison Schools

Type	School	Level
TEAM Comparison for Original TAP school	Green Elementary	Elementary
TEAM Comparison for Original TAP school	Mooreland Heights Elementary	Elementary
TEAM Comparison for Original TAP school	Bearden Middle	Middle
TEAM Comparison for Original TAP school	Cedar Bluff Middle	Middle
TEAM School	Beaumont Elementary	Elementary
TEAM School	Christenberry Elementary	Elementary
TEAM School	Inskip Elementary	Elementary
TEAM School	Maynard Elementary	Elementary
TEAM School	Mount Olive Elementary	Elementary
TEAM School	Norwood Elementary	Elementary
TEAM School	Pleasant Ridge Elementary	Elementary
TEAM School	South Knoxville Elementary	Elementary
TEAM School	Central High	High
TEAM School	Fulton High	High
TEAM School	West High	High
TEAM School	Gresham Middle	Middle
TEAM School	Halls Middle	Middle
TEAM School	Whittle Springs Middle	Middle

Comparisons were made between the pool of TAP and TEAM schools in key performance indicators. Observation scores were baselined by subtracting the school-level average apprentice teacher score, in order to help alleviate numerical issues with varying baselines in observation scores.

Results: Increase in Effective Instruction for Apprentice Teachers

Previous work by the Office of Accountability has indicated that there is learning curve associated with a teacher's classroom observations. The mean observation scores for teachers exhibit a systematic increase until their fifth year of teaching experience. After five years, the mean observation scores exhibit smaller non-systematic drifts.

A comparison of the relative increase in mean observation score by years of experience is included in the Figure 6.1 below. The data comes from three separate years (SY1112, SY1213 and SY1314) and only contrasts the schools that joined TAP in SY1112 with their respective TEAM comparison schools. The earlier TAP schools are not included in this and some of the subsequent analyses. We want to factor out any cumulative effects participation in the TAP program may have on year-over-year performance increases for teachers and students.

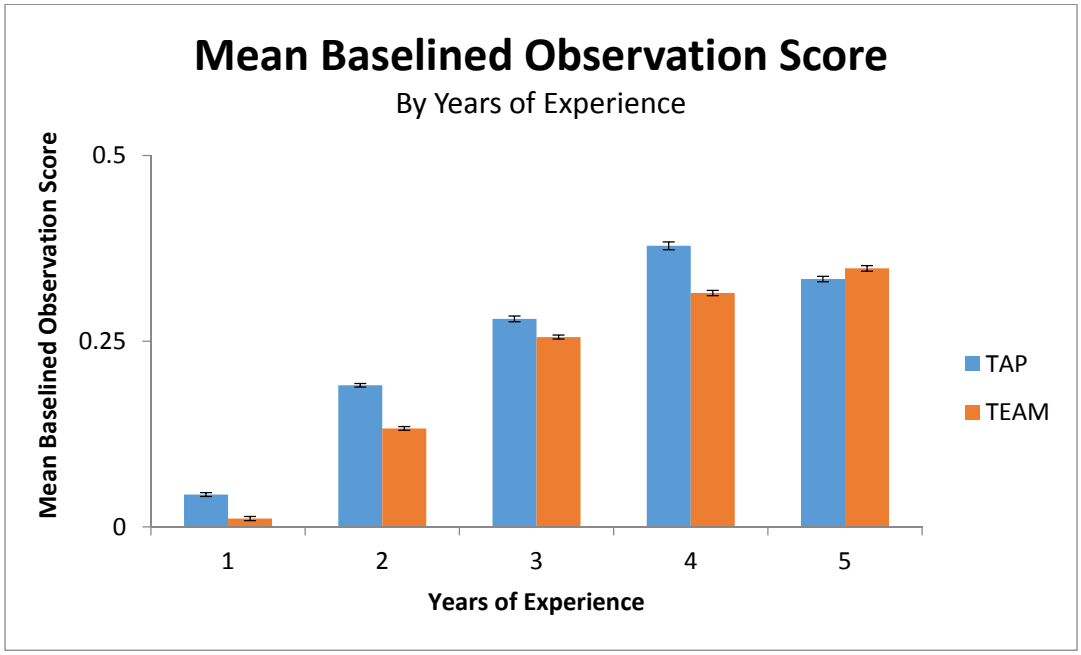


Figure 6.1: Observation Scores by Years of Experience

The mean baselined observation score was statistically higher in TAP schools than in TEAM schools for years 1-4. Interviews were conducted with district staff members who have had experiences with both TAP and non-TAP schools. The interviewees felt that the job-embedded professional development of the TAP system provided teachers with a much fuller and richer understanding of the observation rubric. Figure 6.1 above provides some evidence that this is true.

Results: Student Outcomes Increase

The relative gains in student outcomes are estimated by school-level cumulative (all subject) composite TVAAS gains. School-level gains are considered because TAP and TEAM are initiatives with goals of driving systemic changes at the school-level. Longitudinal mean TVAAS gains for both TAP and TEAM elementary and middle schools (grades 4 through 8) are contained in the Figure 6.2 below. The pool of TAP schools only consists of schools that implemented TAP in SY1112. The original TAP schools are not a part of this data set.

The longitudinal trajectories of the mean TVAAS gains are remarkably similar in both TEAM and TAP schools. Both types of school exhibit peak gains in the first year of implementation of their respective models (SY2011-2012). Disaggregating the TVAAS gains by grade level (elementary versus middle) yields similar results.

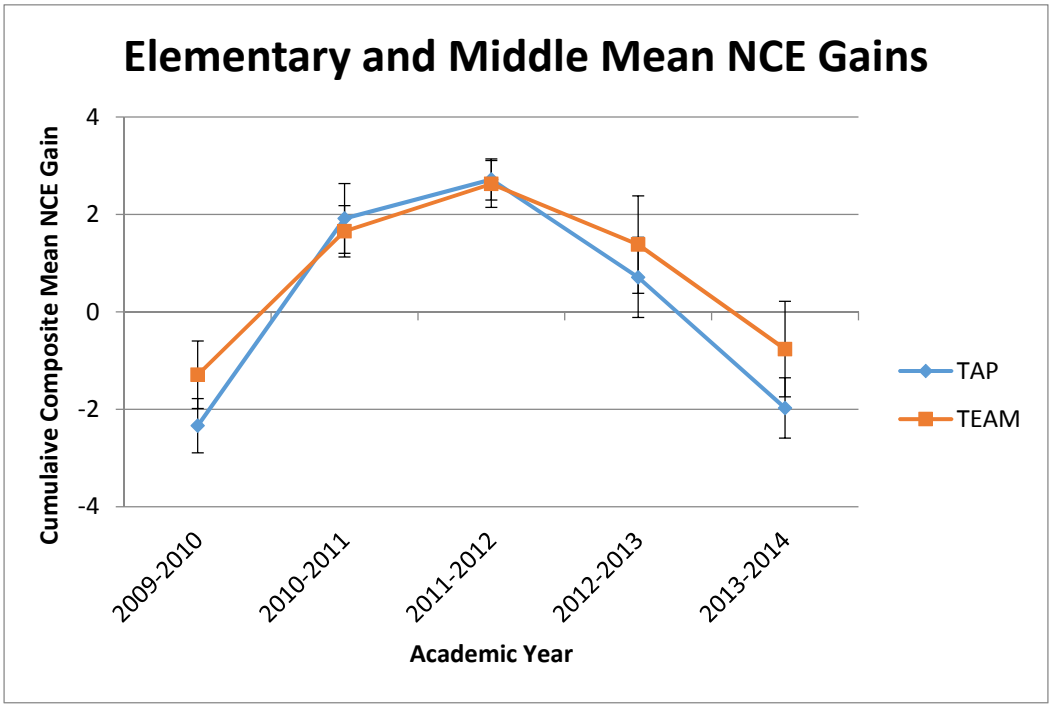


Figure 6.2: Elementary and Middle School TVAAS Trends

The same sort of data in high schools, however, exhibits a diverging trend. Mean TVAAS gains were determined from Algebra I, Biology I, English I, English II, and U.S. History end-of-course (EOC) assessments. There is evidence that the mean EOC gains in TAP schools are higher than the mean gains in the TEAM comparison schools. The point of divergence seems to originate in the year after initial implementation of TEAM and TAP. The results are presented in the Figure 6.3 below.

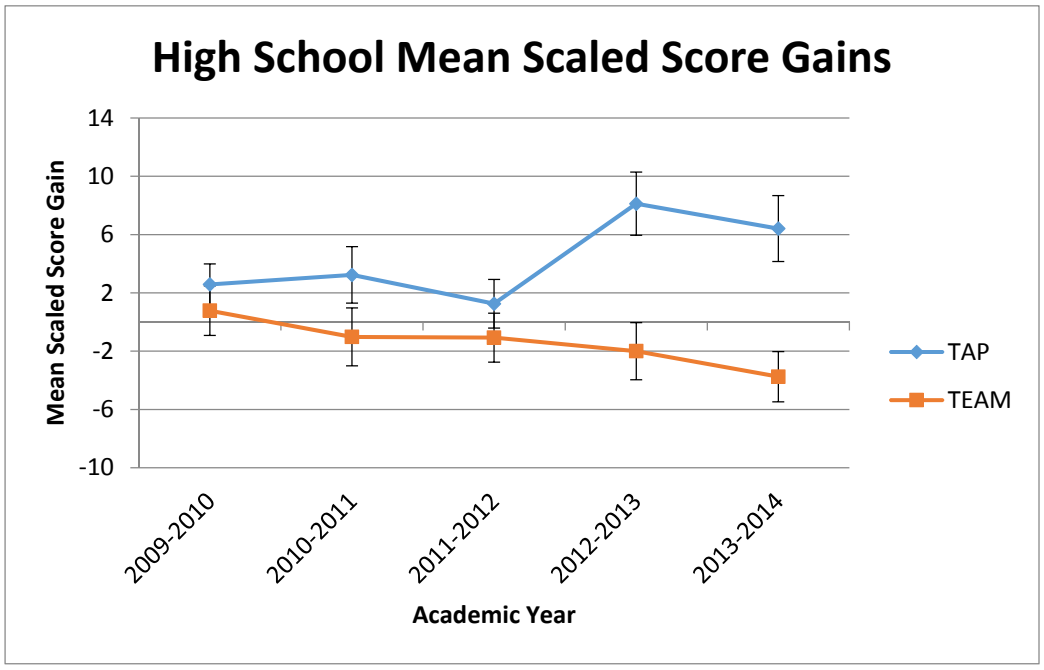


Figure 6.3: High School TVAAS Trends

The data are further disaggregated to the school level to determine if the gains in one TAP school are disproportionally affecting the aggregate trend. Each of the TAP high schools exhibited increases in the mean scaled score gains since implementation. There was a slight dip at implementation, but whereas the TEAM schools have continued on a downward trend, the TAP high schools have improved from their pre-SY11 performance. The same cannot be said for the TEAM comparison schools. Each of these schools exhibited flat or declining mean gains. The trends for the TAP schools can be seen in Figure 6.4 below.

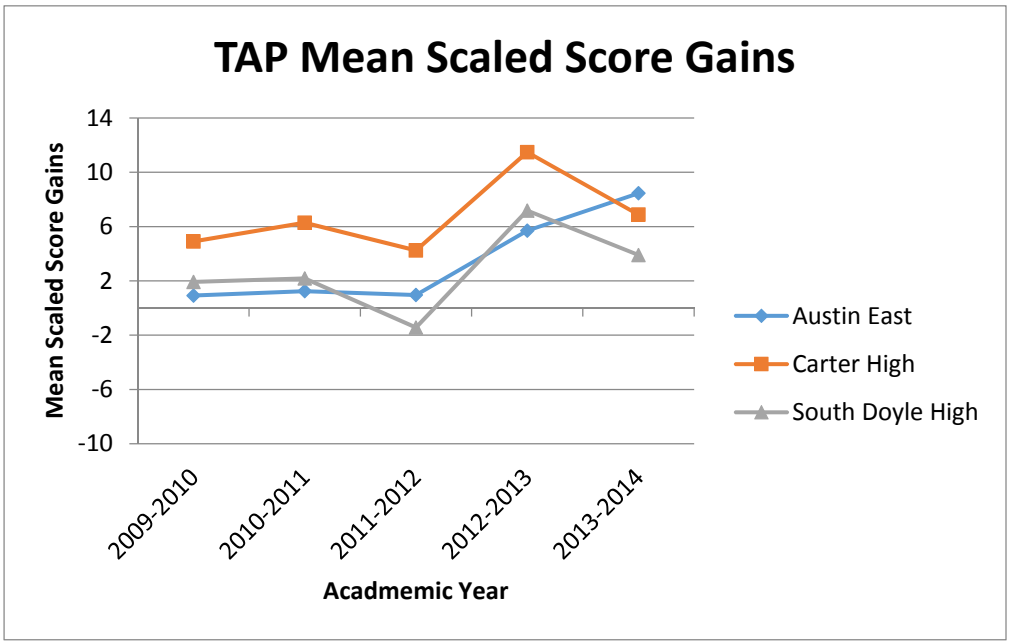


Figure 6.4: Disaggregated High School TVAAS Trends

Comparison of the original TAP schools to the new TAP schools suggests that the new TAP schools experienced a slight directional increase in mean NCE gains in their first year of implementation. This increase was not statistically significant.

This evidence corroborates early findings that the first year of TAP implementation yielded a slight increase in student outcomes.

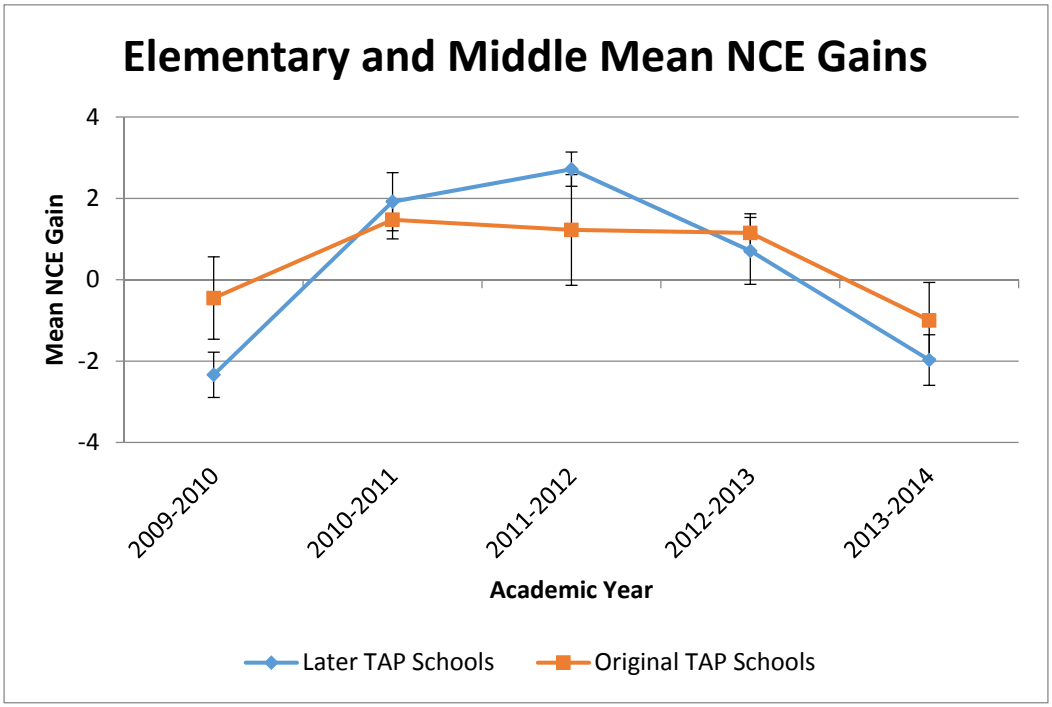


Figure 6.5: TVAAS Gains: Original TAP Schools versus Later TAP Schools

Results: Teacher Perceptions

The survey data referenced in this section was collected anonymously by the New Teacher Center through a contract with the Tennessee Department of Education. Data were collected through the TELL (Teaching, Empowering, Leading and Learning) survey in SY1011 and SY1213. The analysis was completed on a subset of survey questions that aligned with the TAP theory of action. Only the schools implementing TAP in SY1112 and their TEAM comparison schools were included in this analysis. Please note that some schools had too low of participation rates in the questionnaire to be included. The schools included in the dataset are contained in Table 6.3 below.

Table 6.3: TAP and Comparison Schools with Perception Data

Type	School	Level
TAP School	Dogwood Elementary	Elementary
TAP School	East Knox County Elementary	Elementary
TAP School	West View Elementary	Elementary
TAP School	Carter Middle	Middle
TAP School	Austin-East High	High
TAP School	South-Doyle High	High
TEAM School	Beaumont Elementary	Elementary
TEAM School	Christenberry Elementary	Elementary

Type	School	Level
TEAM School	Inskip Elementary	Elementary
TEAM School	South Knoxville Elementary	Elementary
TEAM School	Gresham Middle	Middle
TEAM School	Whittle Springs Middle	Middle
TEAM School	Central High	High
TEAM School	Fulton High	High

The pooled responses for schools in SY1213 are contained below. The TAP schools had responses from 266 of 368 eligible teachers (72.3% response rate). The TEAM comparison schools had responses from 326 of 468 eligible teachers (69.7% response rate). Responses from all of the schools listed in the Table 6.3 above table are contained in Figure 6.6 below.

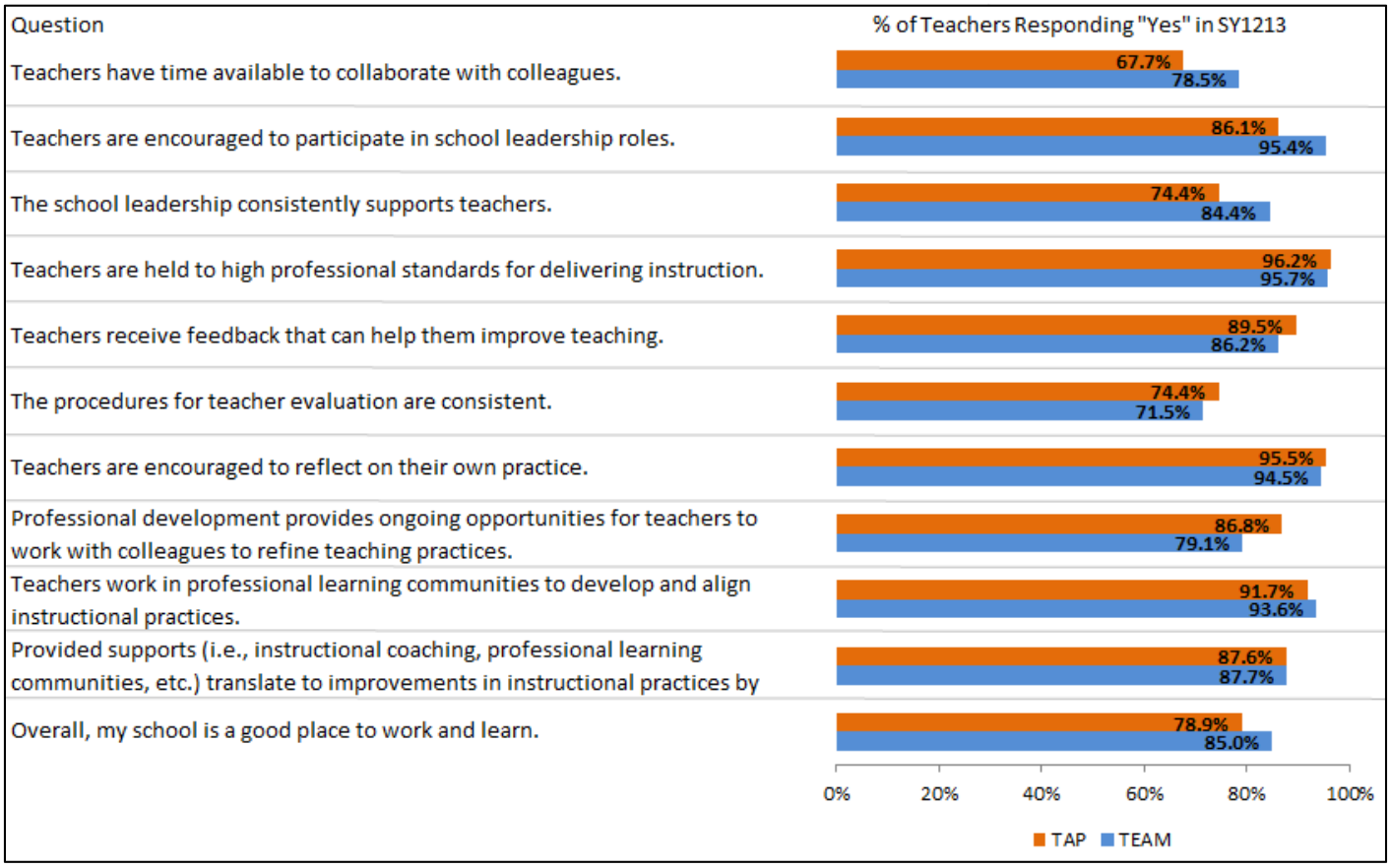


Figure 6.6: TELL Responses

TAP respondents were less likely to feel supported by their school leadership, but had higher positive responses in the other key questions in the school leadership domain. TAP respondents were more likely to answer favorably on questions in the professional development domain. The responses also suggest that TAP respondents were more receptive to feedback delivered through the observation rubric, which confirms information gathered through interviews with TEAM and TAP participants.

Interestingly, a smaller percentage of TAP respondents felt they had available time to collaborate with colleagues, but noted an increased ability to work with colleagues through professional development. This is concerning because the intent of the cluster meetings in the TAP schools was to create time for collaboration. It is possible that respondents prefer peer-to-peer collaboration, but that cluster meetings are less peer-driven and driven more by school leadership.

We also considered the change in responses from SY1011 to SY1213. These data samples suggest the shifts in perceptions that occurred pre-implementation (of both TEAM and TAP) and post implementation. Results are contained in Figure 6.7 below.

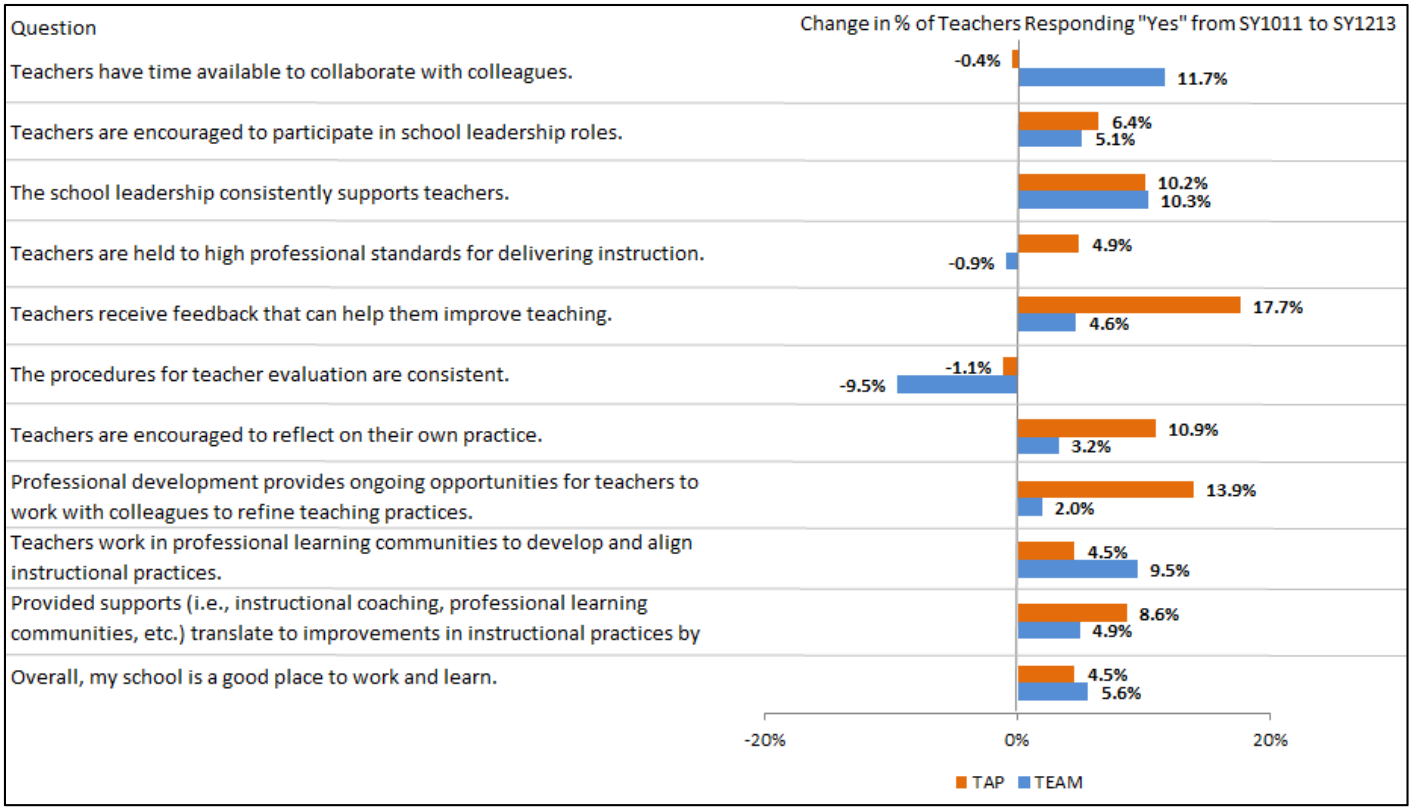


Figure 6.7: Change in TELL Responses: All Schools

There are interesting trends in the data with responses that can be linked to the observation process. TAP respondents had a less dramatic shift in the percentage of teachers that felt the evaluation system was consistent as the transition to the TEAM and TAP systems occurred. TAP respondents were more likely to recognize feedback as keys to improving their practice. This is likely an artifact from the TAP selection process. Teachers had to opt into the TAP system, so movement to more frequent classroom observations was a deliberate choice. In comparison, the TEAM schools were forced to adopt the classroom observations as part of state mandated reforms that stemmed from the Tennessee Race to the Top Grant.

The trends from the TAP schools indicate that fewer teachers felt they had time to collaborate with their colleagues after the implementation of TAP in comparison to pre-implementation. This is a troubling perception, as collaboration is a key tenant of the TAP model. It is possible that relationships between classroom teachers and master/mentor teachers are viewed more as coaching sessions than true collaboration.

The high school TELL survey data are analyzed separately in Figure 6.8 below. The high schools data are pulled out of the main dataset for further analysis because of the divergent trends in TAP and TEAM mean TVAAS gains. The noteworthy results are evident in the data analyzing the change in response rates from SY1011 to SY1213. High school results are contained below.

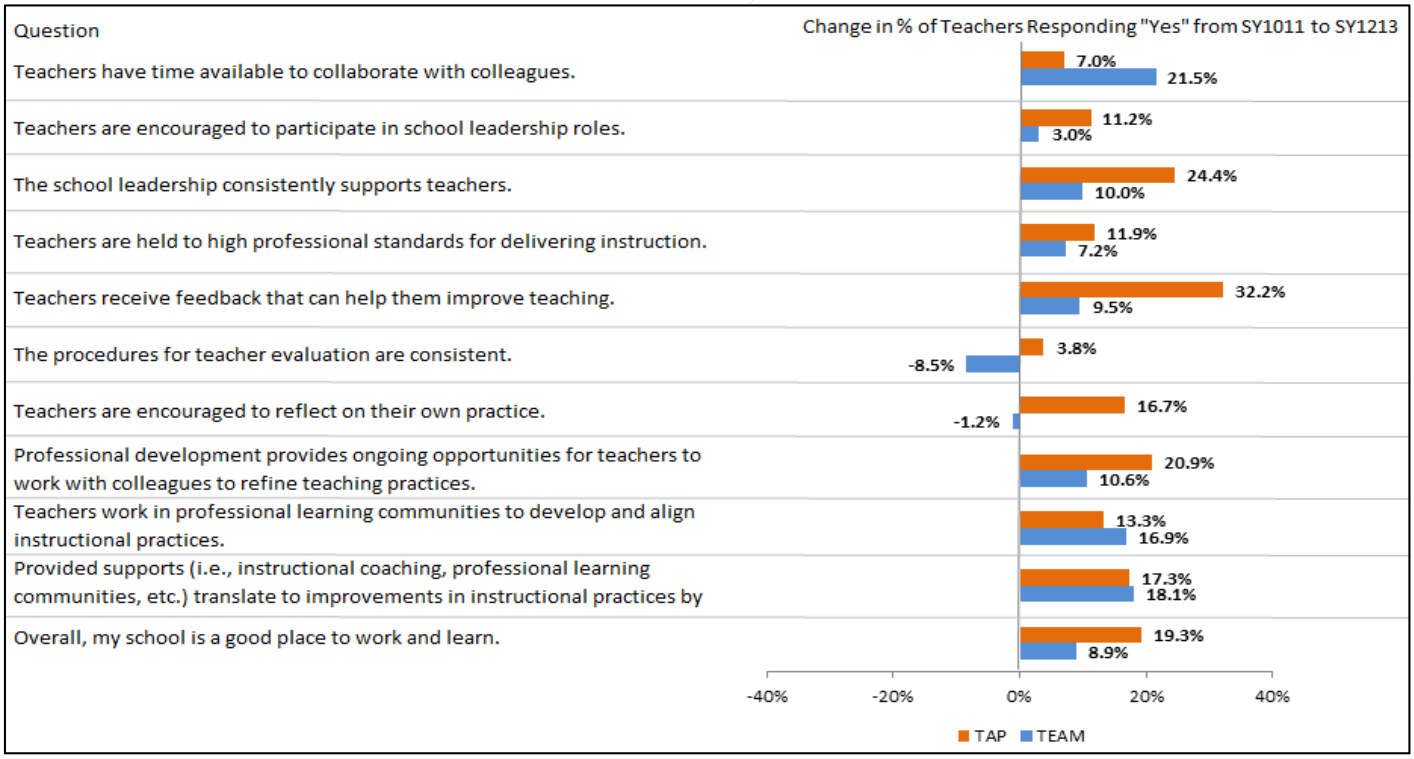


Figure 6.8: Change in TELL Responses: High Schools

What is notable in this data is the magnitude of the changes in responses when considering the dataset as a whole. Consider the responses regarding feedback as a driver for improving teaching. Prior to TAP implementation, 56.8% of respondents at TAP schools felt that they were receiving feedback to improve their teaching. This number increased to 90% after TAP implementation. The magnitudes of the changes were higher among high school TAP respondents when compared to their middle and elementary peers in many key areas. Perhaps this more radical shift is the driving force behind the school-level TVAAS gains.

Results: Teacher Retention

One year teacher retention data was pooled for the SY1112 to SY1213 and SY1213 to SY1314 (the two years with complete retention data since TEAM/TAP was implemented). Retention is defined as staying within the same school from one school year to the next. The data are disaggregated by adjusted summative scores. Adjusted summative scores include individual TVAAS scores for tested teachers, but omit growth data for non-tested teachers. The dataset excludes the original TAP schools and their respective TEAM comparison schools. Results are contained below.

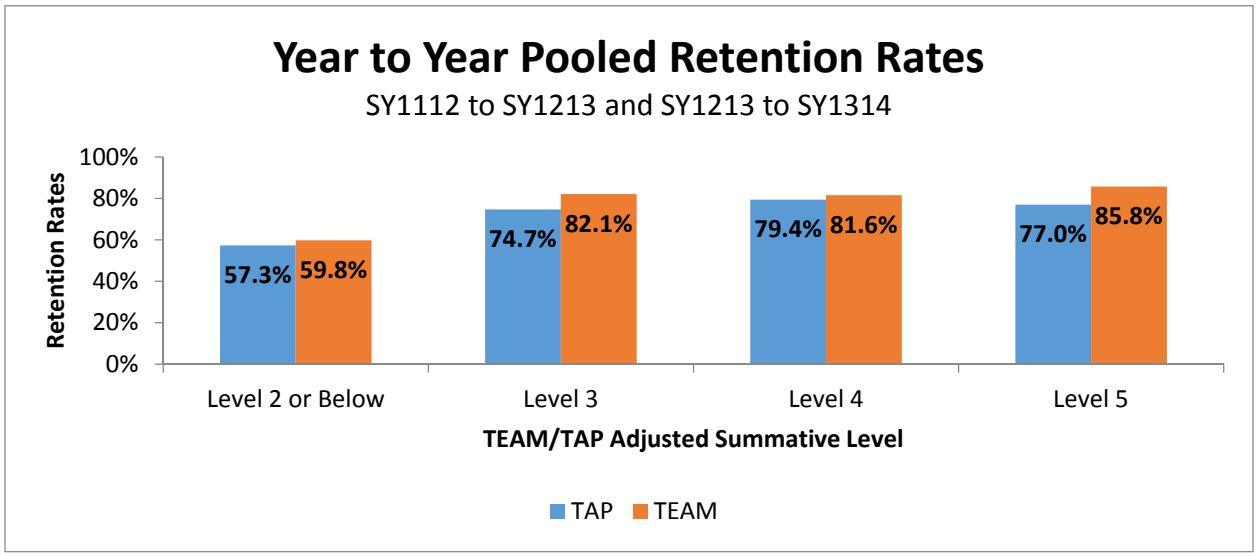


Figure 6.9: Year to Year Teacher Retention Rates

Chi-squared testing indicates there is no statistical difference in the teacher retention distributions (by adjusted summative level) for TEAM and TAP schools ($p = 0.131$). Additional chi-squared testing indicates that the distribution of retained teachers who were level 3 or above is not the same in TEAM and TAP schools ($p= 0.000$). Visual inspection of the distribution indicated that TEAM schools have a higher rate of retention of teachers who scored level 3 or higher on their respective rubrics. There is no statistical difference in the distribution of teachers who were level 2 or below ($p = 0.462$).

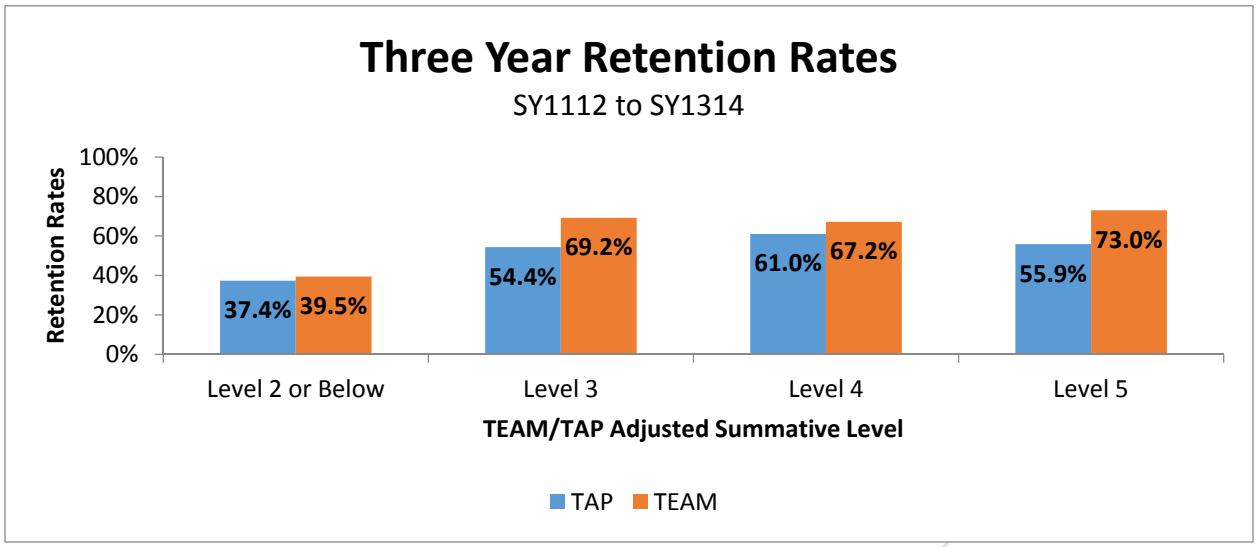


Figure 6.10: Three Year Teacher Retention Rates

Chi-squared testing indicates that the distribtuion of three year retention rates by adjusted summative level are not the same in TEAM and TAP schools (p = 0.005). Further testing indicates that the retention rate of teachers who score level 2 or below is no different between TEAM and TAP schools (p = 0.668). However, the retention rate of teachers rated 3 and above is not the same (p = 0.000). Inspection indicates that TEAM schools have higher retention rates for teachers who score level 3 or higher.

Turnover may not be such a bad thing in TAP high schools. Anecdotally, some of the high schools are using TAP supports, professional development and monetary bonuses as a recruiting tool. There is some evidence to suggest that TAP high schools may be recruiting higher potential cadidates from the general hiring pool. Analysis of the TVAAS gains (in SY1112, SY1213 and SY1314) of high school teachers by years of experience is contained below. There were insufficent N counts in years of experience bands beyond three years to make the data meaningful.

Table 6.4: TVAAS Gains by Years of Experience: High Schools

Years of Experience	Type	N	Mean TVAAS Gain	Std. Deviation	p-value
1	TAP	19	6.53	8.21	0.001
	TEAM	29	-3.14	10.15	
2	TAP	26	1.42	10.67	0.123
	TEAM	35	5.43	9.27	
3	TAP	18	6.94	9.57	0.601
	TEAM	31	5.32	10.85	

The mean gain for first year teachers was statistically different in TAP and TEAM schools. First year TAP teachers out-performed their TEAM peers. Beyond the first year of experience, however, there is no statistical difference in mean TVAAs gain. This pattern is not evident in the elementary and middle school data. However, it is unlikely that increases in just this subset of teachers is large enough to drive the school level TVAAS gains presented earlier.

Conclusions and Considerations

Generally the results of this study mirror results presented in a 2012 NIET study of the TAP system. Some of the salient points of that study and their relationship to this study are contained below.

- According to NIET, TAP schools show consistently high rates of student growth. That study did not specifically compare against a control group to determine the magnitude of relative rates of growth. While there is no evidence in this study to say that TAP is having a measurable impact on school-wide and teacher TVAAS in elementary and middle schools, there is some evidence that high school growth gains are being positively impacted by TAP.
- Per the NIET study, TAP teachers show growth over time in quality of their instruction. There is some evidence that this is occurring in KCS schools as evident in longitudinal teacher observation scores. The TAP teachers in this study exhibited evidence of greater mean growth as measured by the classroom observation rubric. However, the magnitude of these changes may not be enough to impact student outcome data in the elementary and middle schools.
- TAP increases the recruitment and retention of highly effective teachers according to the 2012 NIET study. There is some evidence of better recruitment in all grade bands. Relative mean teacher evaluation scores are higher for teachers early in their career. This could be a result of more effective teaching, but it also could be the result of the level of support and professional development that occurs around the rubric in TAP schools. However, there is no evidence in elementary and middle schools that this increase in observation scores translates to gains in student outcome data. There is some evidence to suggest that TAP may be having an impact on student level data at the high school level, and further evidence that high schools may be using TAP as a tool to recruit teachers that are more effective early in their career. There is little evidence to support that TAP is a driver for teacher retention. Retention rates of higher performing teachers are better at TEAM schools.
- Stakeholders report high levels of support for professional development in TAP schools in the NIET study. This echoes some of the survey data. TAP schools

outperformed their TEAM comparison schools on positive response rates around questions relating to professional development and professional learning. This may be evidence of the effectiveness of TAP field testing, clusters and associated "follow-up."

Whereas this analysis helps confirm some of the environmental and rubric-focused conclusions of the NIET study, the student outcome results appeared to be mixed. There is evidence in middle and elementary TAP schools that teachers feel they have access to support and quality professional development. However, the difference between TEAM and TAP may not be significant enough to affect the middle and elementary student outcome data.

There is evidence that student outcome data at TAP high schools is out-pacing student outcome data at TEAM schools. However, district TAP experts are at a loss as to explaining why the impacts of TAP on student gains would be different at high schools. What is even more perplexing is that TAP External reviews tend to give the lowest quantitative scores and the less favorable qualitative scores to the high schools. It is possible that the student gains are impacted by an unobserved variable unrelated to TAP. As such, this analysis raises some additional research questions that may be the basis for future studies. Most notably, what is being done differently in TAP or TEAM high schools to yield such divergent student outcome data? A more focused study around high school TAP implementation may be warranted. The work presented in this analysis is insufficient to create a causal link between teacher perceptions, student outcomes, and TAP/TEAM implementation.

Future budgeting concerns may make it impossible to continue TAP in all of the currently participating schools. There may be insufficient resources in the general purpose budget to absorb all of the costs associated with TAP as the grants that fund the TAP initiative come to the end of their life cycles. This represents an opportunity to study the effect of the removal of the TAP programs from schools. Two different studies are proposed below. The study chosen will be dictated by the research question that is expected to be answered.

1. Schools that no longer wish to be a part of TAP can be removed from the program. The schools that are interested in remaining in the TAP program can apply for a random lottery. Schools that win the lottery will continue with the TAP initiative.
2. A targeted withdraw can be done to satisfy budget requirements. Preference for continuation should be granted to high schools (who are demonstrating the student gains) and priority/focus schools to retain the supports that are already in place.

APEX

In SY1112, Knox County Schools (KCS) implemented a strategic compensation system called Advance, Perform, Excel (APEX). The APEX compensation system was designed over a one year timeframe with broad input from a cross-functional team. Participants in the design of APEX included KCS staff in human resources, finance, and accountability; teachers (via surveys, focus groups and town-hall style input sessions); principals (via focus groups); and the Knox County community (via community forums). The theory of action of APEX is to reward high quality instruction, student achievement, teacher leadership, and continued service in high-needs schools.

Under the APEX compensation system, eligible staff members can earn up to $2,000 per year in bonus pay. The pertinent metrics that serve as inputs to the APEX calculations include TEAM observation data, a variety of student outcome metrics, leadership rubric scores and years of continual service in the same high-needs schools.

It is important to note that TEAM and APEX were launched in the same year. It will be impossible to create any causal links between outcome measures and APEX because of how intertwined it is with the TEAM process.

Methodology

The analysis uses a mixed methods approach to estimate the impact and perceptions surrounding APEX. Survey data regarding the implementation of the strategic compensation programs (APEX and TAP payouts) are collected annually through the Tennessee Consortium on Research, Evaluation, and Development (TNCRED). The Tennessee Value Added Assessment System (TVAAS) provides longitudinal data regarding student performance. Other data sources used in this analysis include the KCS human resources database and the district level repository of observation data for certified staff (RANDA Tower).

Results: TNCRED Data Trends – Perceptions about Program Implementation

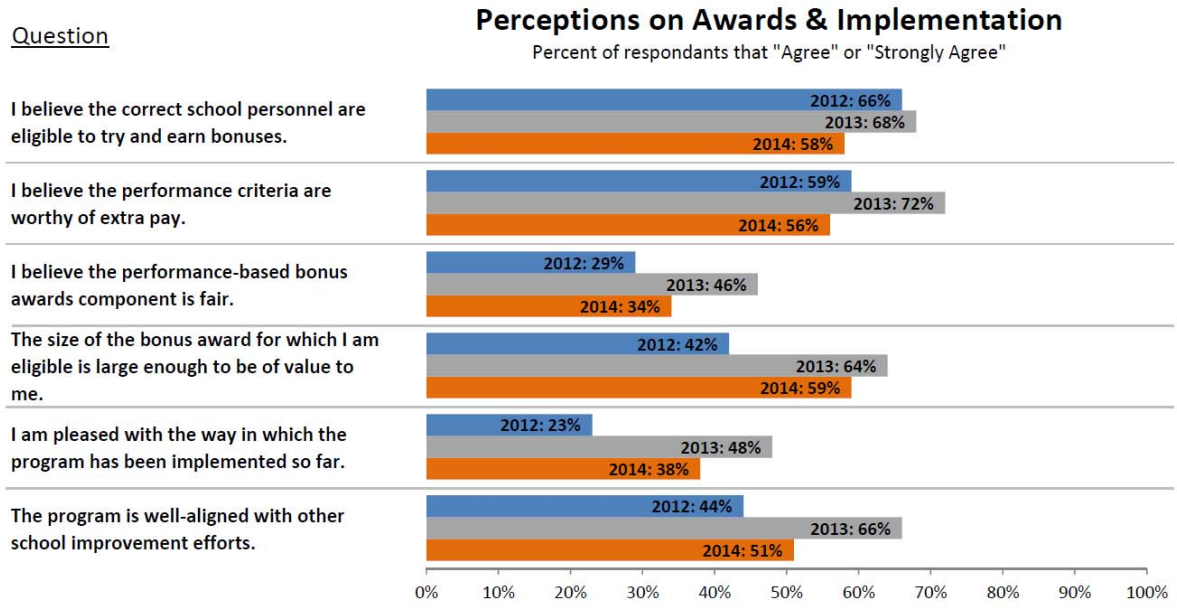


Figure 7.1: TNCRED Perception Data: Award and Implementation

The responses from key TNCRED survey questions are contained below. Survey responses are aggregated by year (starting with SY1112). The number of respondents by year are SY1112=2174, SY1213=2547 and SY1314=1932. Please note, the data for SY1314 are still considered preliminary. The survey respondents include teachers that are not eligible for APEX, as they are paid bonuses through the TAP system. TNCRED did not disaggregate TAP and TEAM respondents in SY1112 and SY1213. The disaggregation of SY1314 data will help shed light on the diverging opinions of TAP and TEAM staff. Further note that the survey data are collected before the award notification (for that academic year) has been made. For example, teachers completed the SY1314 survey prior to the release of the student outcome data that is required to calculate their SY1314 APEX/TAP awards.

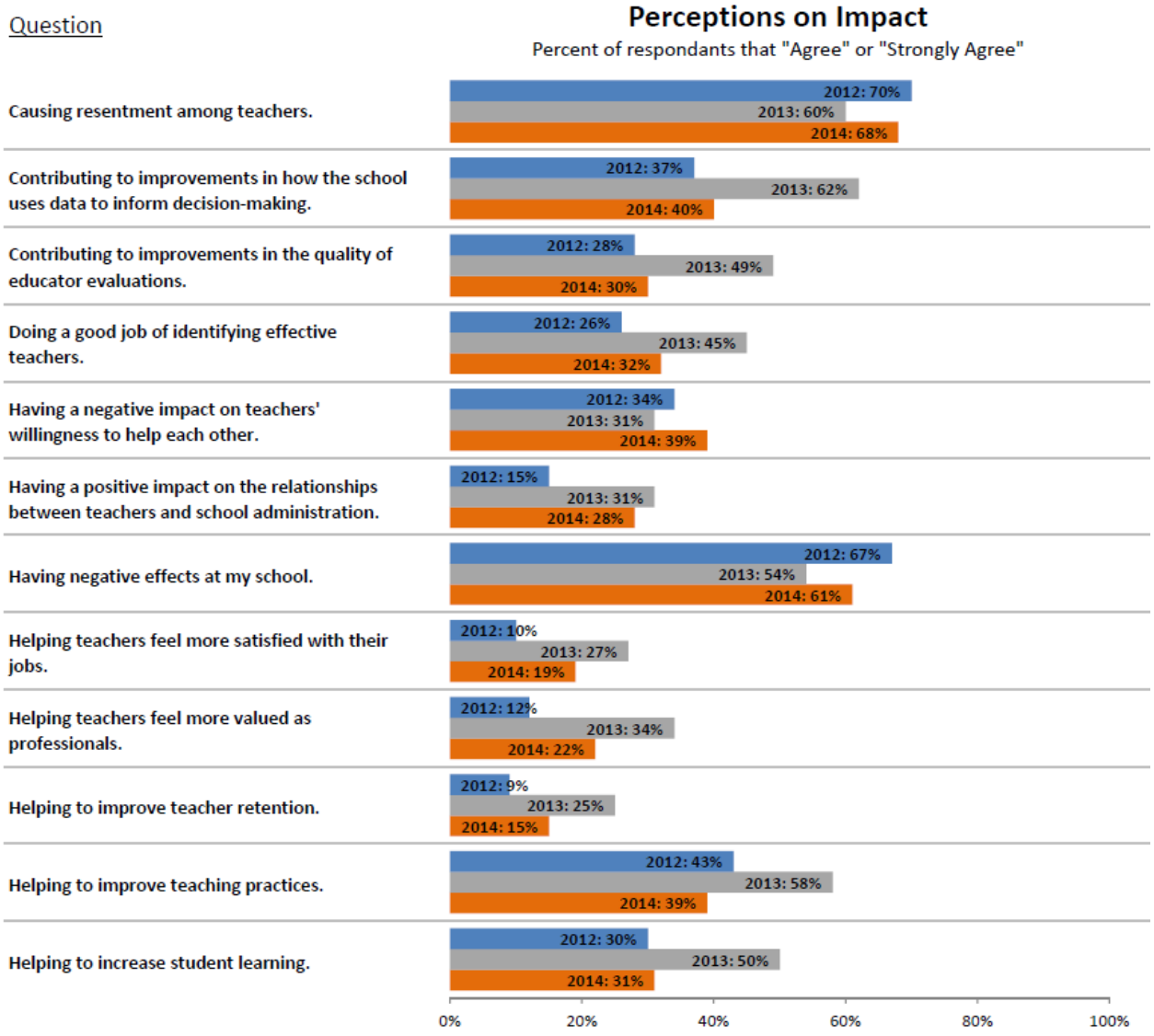


Figure 7.2: TNCRED Perception Data: Impact

It is interesting to note that SY1213 represented a high point in terms of positive responses regarding the strategic compensation program. This survey was taken after the highest year in terms of the percentage of staff receiving APEX payouts (SY1112). Figure 7.3 below shows the trends in the percent of teachers receiving APEX.

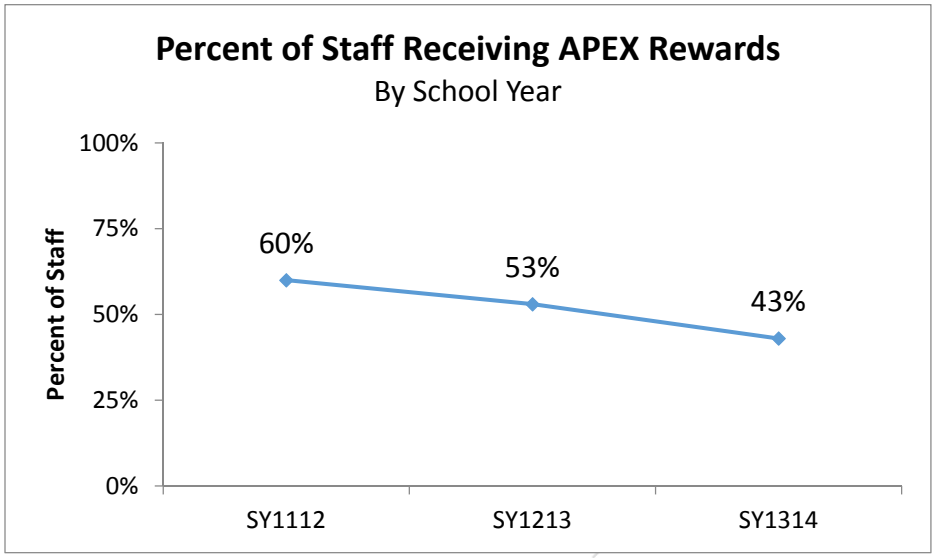


Figure 7.3: APEX Payout Rates

A majority of respondents in each year felt that the strategic compensation causes resentment among teachers. Similarly, a majority of respondents disagreed (in each year) that the models do a good job of identifying effective teachers. A very small portion of respondents agreed that the strategic compensation program helps teachers feel more valued as professionals. Some of the key reasons why APEX exists are to reward effective teachers and to help them feel more valued.

The SY1314 data are disaggregated by the staff member's observation system (TEAM vs. TAP) so that we can estimate the effect of including TAP teachers in the survey data. Results are contained in Table 7.1 below. The p value represents the probability that difference in the distribution of respondents that agree or strongly agree with the statement is no different between TAP and TEAM teachers.

Table 7.1: Perception Data: TEAM Vs. TAP

% of Respondents that Agree or Strongly Agree in SY 1314			
Question	**TEAM**	**TAP**	**p-value**
I believe the performance-based bonus awards component is fair.	27%	55%	3.4E-82
I believe the performance criteria are worthy of extra pay.	50%	75%	3.6E-52
I believe the correct school personnel are eligible to try and earn bonuses.	55%	66%	1.5E-11
I am pleased with the way in which the program has been implemented so far.	32%	61%	2.4E-108
The program is well-aligned with other school improvement efforts.	45%	72%	3.7E-83
The program is well-aligned with other FTTT initiatives at my school.	51%	73%	4.6E-55
I feel that the program has been responsive to teacher feedback and needs.	22%	44%	3.0E-79
The program has added burdensome paperwork to teachers' workload.	50%	56%	1.9E-05
The program requires completing tasks that take time away from planning and instruction.	49%	57%	9.8E-09
The TDOE can adequately support the program.	33%	47%	3.8E-26
My district's central office can adequately support the program.	33%	47%	3.8E-26
I believe the financial resources exist to sustain the program over time.	21%	36%	3.0E-39
I believe the program has the support of the local community.	31%	47%	5.5E-35
Having negative effects at my school.	64%	44%	1.2E-49
Causing resentment among teachers.	71%	59%	4.7E-21
Having a negative impact on teachers' willingness to help each other.	42%	28%	3.9E-24
Helping teachers feel more satisfied with their jobs.	15%	33%	5.6E-72
Helping teachers feel more valued as professionals.	17%	38%	3.3E-88
Contributing to improvements in the quality of PD offered to teachers.	32%	50%	6.0E-43
Helping to improve teaching practices.	32%	61%	2.4E-108
Helping to increase student learning.	25%	56%	1.2E-142
Helping to improve teacher retention.	12%	31%	1.2E-95
Contributing to improvements in how the school uses data to inform decision-making.	34%	62%	1.6E-98
Contributing to improvements in the quality of educator evaluations.	24%	52%	5.4E-120
Having a positive impact on the relationships between teachers and school administration.	16%	35%	1.8E-75
Having positive effects at my school.	22%	51%	1.1E-136

TAP teachers agreed or strongly agreed more frequently with all the positive statements regarding strategic compensation and agreed or strongly agreed with far fewer negative statements regarding strategic compensation in SY1314. This provides some evidence that

inclusion of the TAP teachers in the TNCRED data is serving to inflate agreement with positive statements and disagreement with negative statements.

Results: Teacher Leadership

Certified staff are rated on the amount of leadership responsibilities they assume as part of the KCS observation process. These leadership scores are used as components of the APEX calculation. The distributions of leadership scores by academic year (for all APEX eligible positions) are contained below. N counts are as follows: SY1112=3464, SY1213=3632, and SY1314=3609. P values represent the probabilities that the distribution of teachers within certain categories is no different when comparing the baseline year to the comparison year.

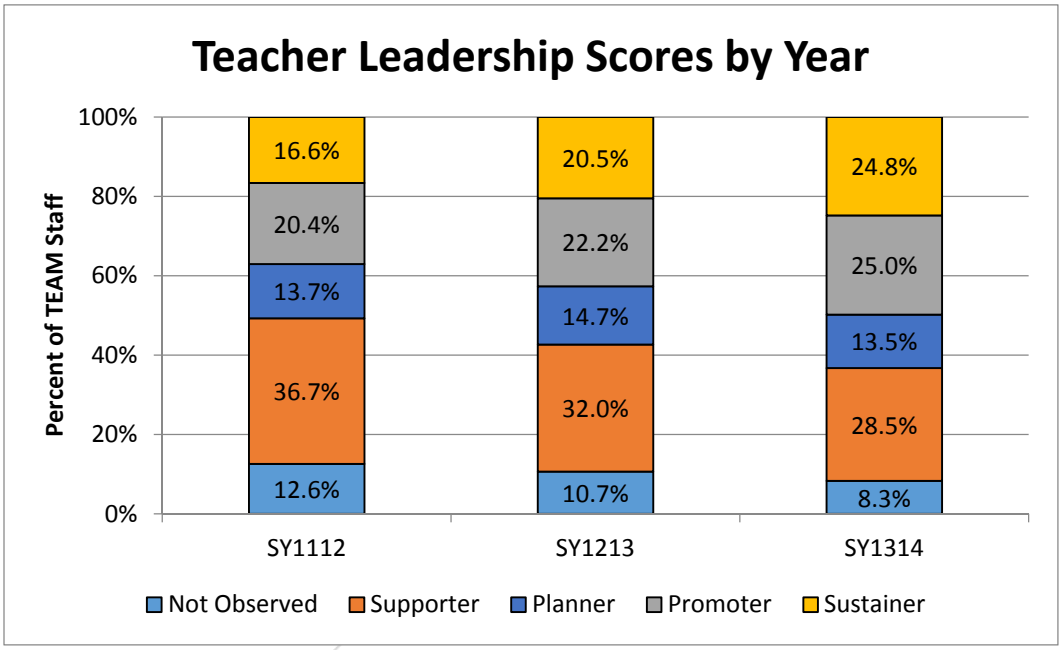


Figure 7.4: Teacher Leadership Distributions

Table 7.2: Non-parametric Test Results on Leadership Distributions

Group	Baseline Year	Comparison Year	p value
Distribution of staff scoring a 0 on the leadership rubric	SY1112	SY1213	0.00071*
	SY1112	SY1314	0.00000*
	SY1213	SY1314	0.00001*
Distribution of staff scoring a 20 on the leadership rubric	SY1112	SY1213	0.00000*
	SY1112	SY1314	0.00000*
	SY1213	SY1314	0.00000*

Inspection of the data indicates that the distribution of staff members earning no leadership points has decreased each year under the APEX system. It also indicates that the percentage of teachers earning the highest possible ratings on the leadership rubric has increased year over year. This is evidence that the number of teachers assuming leadership responsibilities has increased since the implementation of APEX.

Results: Student Outcomes Increase

Piecewise analysis of the longitudinal changes in school-wide EOC and TCAP TVAAS are contained below. Data is only included for subjects that have had state TCAP and EOC assessments since SY0910. The subjects included are Algebra I, English I, English II, Biology I and US History. TCAP data is limited to grades 3-8. Figure 7.5 below displays TVAAS gains based on TCAP and EOC results. Please note: the data displayed below is grades 4-12 since KCS early grades (including third grade) did not always generate value-added scores.

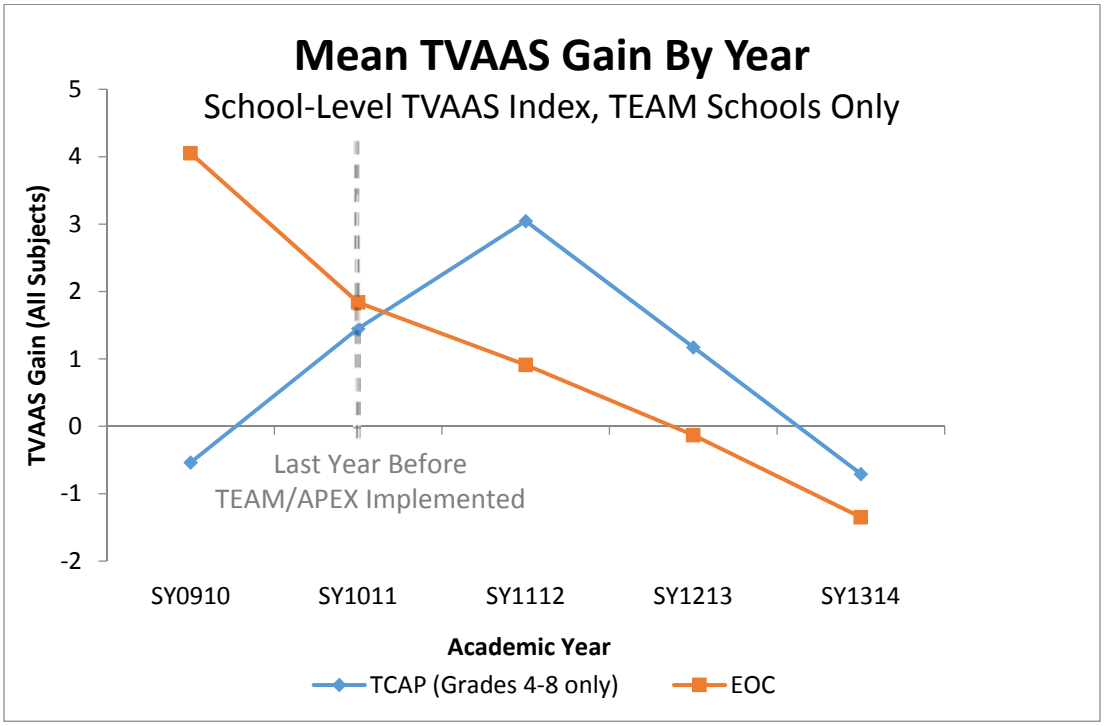


Figure 7.5: Mean TVAAS Gains by Year

The slope of the line through the SY0910 and SY1011 EOC gains is -2.214. This means that the mean amount of growth in TEAM high schools was decreasing at a rate of 2.14 points per year. There is evidence that the rate of loss in EOC scaled score points slowed after the implementation of TEAM/APEX. The slope of the line through data points in SY1011 through SY 1314 is -1.061 which indicates that the mean amount of growth in TEAM high schools is now decreasing at a rate of 1.061 scaled score points per year. We are 95% certain that the

slope after TEAM/APEX implementation is between -1.259 and -0.862. Since this range does not include the original slope of -2.214, we are relatively confident that the trajectory of scaled score growth in high schools is different after implementation of TEAM/APEX.

The TCAP data is more difficult to analyze without a direct (intra-year) control group. A dramatic change in trajectory of TCAP gains occurred after the end of SY1112. This corresponds to the second year of TEAM/APEX implementation. It is important to note that since SY1112, state performance indicators (SPIs) have been systematically narrowed on the TCAP assessment as a precursor to implementing new state standards.

The highest rate of APEX payout occurred in SY1112 when the mean gains on the TCAP were highest. The gradual decrease in the number of teachers earning APEX has decreased as the mean TVAAS gains have decreased.

Results: Increase in Effective Instruction

Teacher observation scores are difficult to compare year to year. Issues arise with this data due to inter-rater reliability and shifting baselines between schools. In addition, the TEAM observation structure changed in SY1213 and SY1314. In SY1213, each observation of professionally licensed staff members was unannounced. This was a change from SY1112, where professionally licensed staff members had one announced and one unannounced observation. Staff members had an option of knowing the month of their final observation in SY1314. The observation scores are therefore compared to an intra-year and intra-building baseline to help mitigate some of the effects above. Adjusted summative observation scores are used in this analysis. Adjusted summative scores include observation scores but will only include student outcome data for teachers generating individual growth scores.

Each school has a different baseline. The average apprentice level score (calculated for each school) is subtracted from each staff member's observation score in order to provide a base-lined measure of teacher effectiveness. A p-value represent the probability that the mean base-lined adjusted summative score is no different from one year to the next. Only veteran teachers are included in the data below.

Table 7.3: Base-lined Teacher Observation Scores by Year

Metric	Years	N	Mean Base-lined Adj. Summative	Std. Deviation	p-value
Difference from Average Apprentice Teacher by Year	SY1112	2667	26.055	55.134	.000*
	SY1213	2649	39.173	57.543	
	SY1213	2649	39.173	57.543	.006*
	SY1314	2650	43.547	59.387	
	SY1112	2667	26.055	55.134	.000*
	SY1314	2650	43.547	59.387	

There is statistically significant evidence that the mean adjusted summative scores for professional teachers has increased each year since TEAM/APEX was implemented.

Results: District-Level Teacher Retention

Teacher retention rates by year are presented in Table 7.4 below. Retention is classified as continuing employment as certified staff in the district from one year to the next. Teachers are permitted to move from one school to the next and still be considered retained. Retention rates are disaggregated by their adjusted summative level of effectiveness. The SY1314 data should be considered preliminary. The final value will be adjusted downward as teachers who were employed in SY1314 continue to leave the district through the duration of SY1415 (through retirement, resignations, and terminations). Professional and apprentice teachers are included.

Table 7.4: Retention Rates by TEAM Adjusted Summative Scores

Year	Retention Rates by Adjusted Summative Level – TEAM schools only							
	N	1	2	3	4	5	Exempt	All Staff
SY1011	3506	-	-	-	-	-	-	89.7%
SY1112	3268	0.0%	80.5%	89.7%	90.6%	90.5%	66.7%	89.8%
SY1213	3468	35.7%	76.3%	87.9%	89.8%	90.4%	25.3%	87.0%
SY1314	3526	68.4%	83.6%	90.5%	93.4%	93.2%	53.7%	90.5%

Chi-squared tests were performed on the distribution of teachers that were retained. Results are contained below. The p value is the probability that the distribution of retained teachers is no different from the baseline year to the comparison year. The "teachers that meet or exceed expectations" category corresponds to teachers with an adjusted summative score of level 3 or higher. The "Teachers that did not meet expectations" category corresponds to teachers who scored level 1 or 2 on their adjusted summative scores. Directional analysis indicates that retention rates were lower in SY1213 than SY1112, but increased in SY1314 when comparing to both SY1112 and SY1213. Results of the statistical analysis of this data are below.

Table 7.5: Changes in Teacher Retention Rates

Category	Baseline Year	Comparison Year	Baseline Retention	Comparison Retention	p-value
Teachers that meet or exceed expectations	SY1112	SY1213	90.3%	89.4%	0.058
	SY1112	SY1314	90.3%	92.4%	0.000*
Teachers that did not meet expectations	SY1112	SY1213	79.2%	73.6%	0.000*
	SY1112	SY1314	79.2%	81.9%	0.000*

The distribution of retained teachers that meet or exceed expectations was no different in SY1213 than it was in SY1112 at the 95% confidence limit, but it is very close. The retention distribution for teachers the meet or exceed expectations is statistically different when comparing SY1112 to SY1314. Visual inspection indicated that SY1314 has a higher retention rate of teachers that exceed expectations. However, the same is true for teachers that did not meet expectations. Visual inspection indicated that retention of teachers that did not meet expectations was higher in SY1314 than it was in SY1112. The converse is true when comparing SY1112 to SY1213.

Results: Retention in High-Needs Schools

A high-needs metric is added to the APEX calculation to provide a boost in the APEX calculation for teachers in traditionally hard to staff schools. There are two tiers of metrics awarded. Schools that serve a population with at least 75% of students eligible for free or reduced price lunch (FRL) are placed in Tier I and receive the highest value to their APEX calculation. Schools that serve a population of students greater than 50% FRL but less than 75% FRL are placed in Tier II and receive a lower value to their APEX calculation. Schools that serve a population of students that are less than 50% FRL are not considered high-needs and receive no additional value to their APEX calculation. Teachers that remain within the same high-needs school (regardless of Tier) receive the maximum value to their APEX score after serving for three consecutive years in the same school. As such, the data below will examine three year retention rates within the same school. Retention data are contained in the Figures 7.5 and 7.6 below.

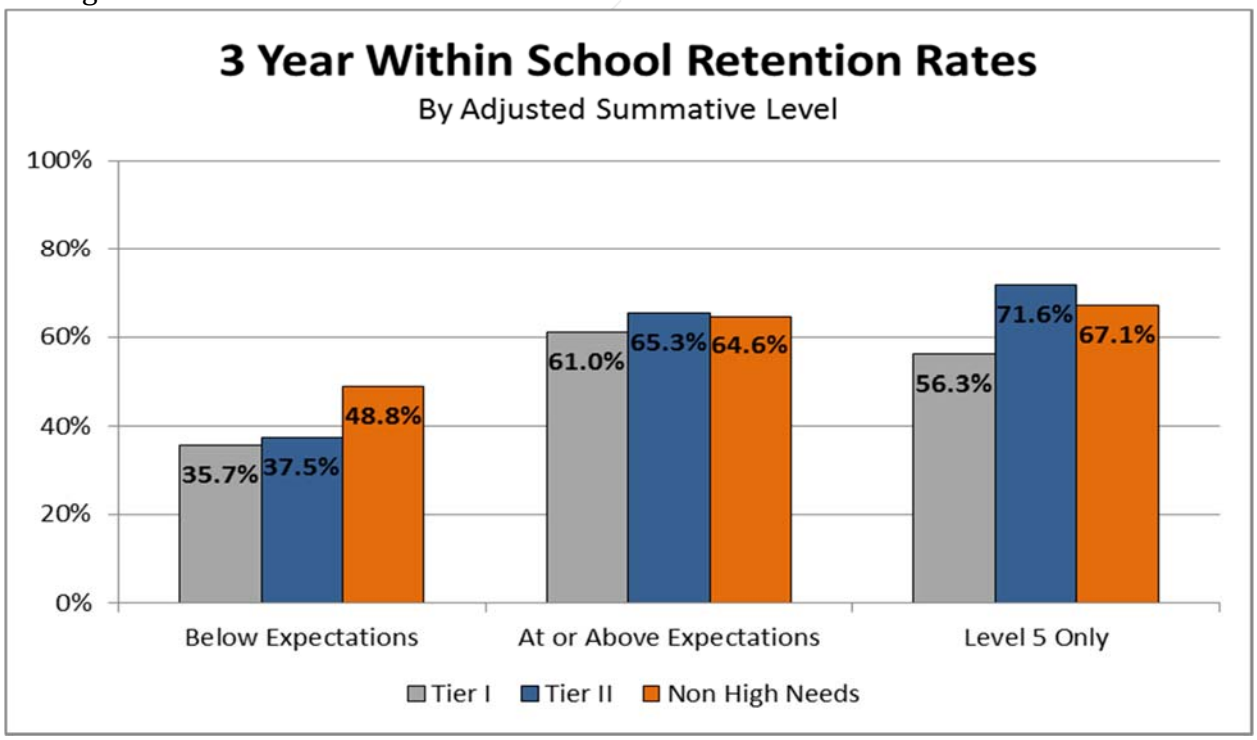


Figure 7.6: Three Year Within-School Retention Rates: By Adjusted Summative

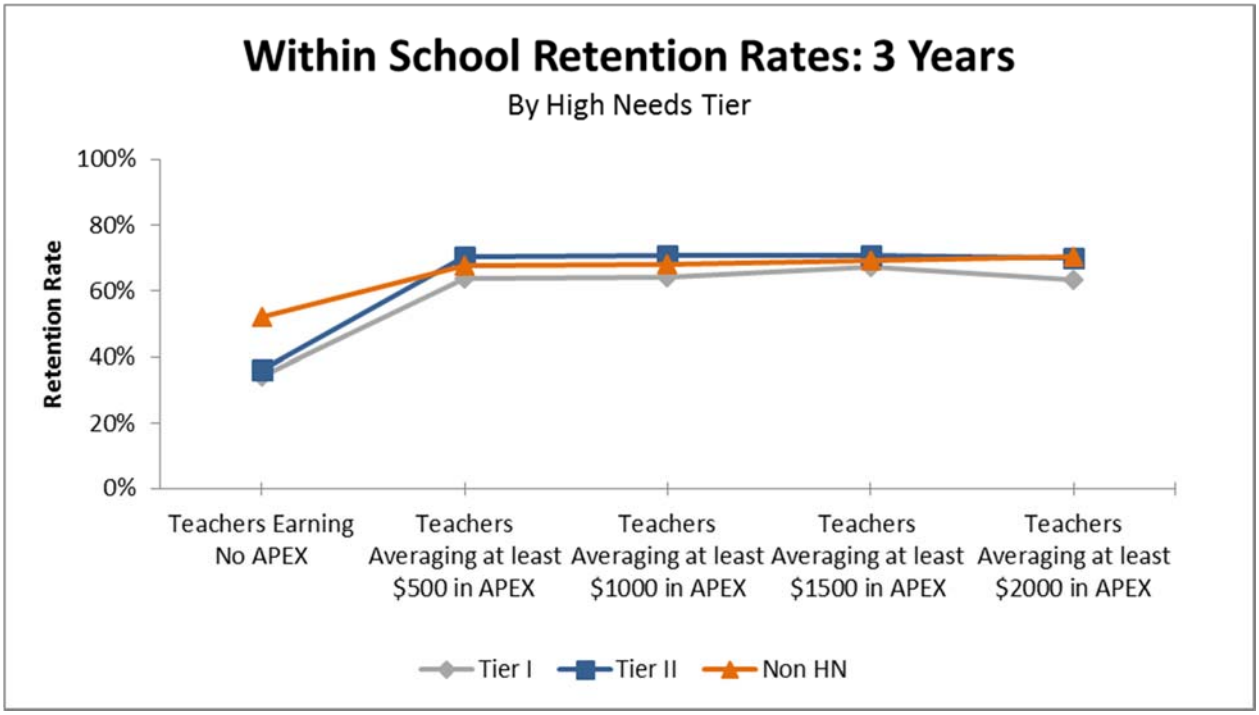


Figure 7.7: Three Year Within-School Retention Rates: By High Needs Tier

Visual inspection suggested that earning any amount of APEX (even as low as averaging $500 per year) may have a dramatic impact on a teacher's retention in high-needs schools. However, respondents to the TNCRED survey perceive that strategic compensation has little impact on teacher retention. To counter this point, TNCRED data is derived from the entire pool of Knox County teachers. KCS conducted a one question survey in order to better estimate the impact of APEX on three year retention rates in high-needs schools. The survey was conducted among the pool of certified staff that have served continuously in the same high-needs school for at least the last three years. The surveyed staff members were asked to rank a list of five factors in their relative importance to their continued service in high-needs schools. The list was in no way intended to be exhaustive, but rather was to provide a relative importance of the five factors. Results from the survey are contained below. The number of respondents serving in a Tier I school ranged from 104 to 114 depending on the factor being scored as some were left unranked. The number of respondents serving in Tier II schools ranged from 115 to 124.

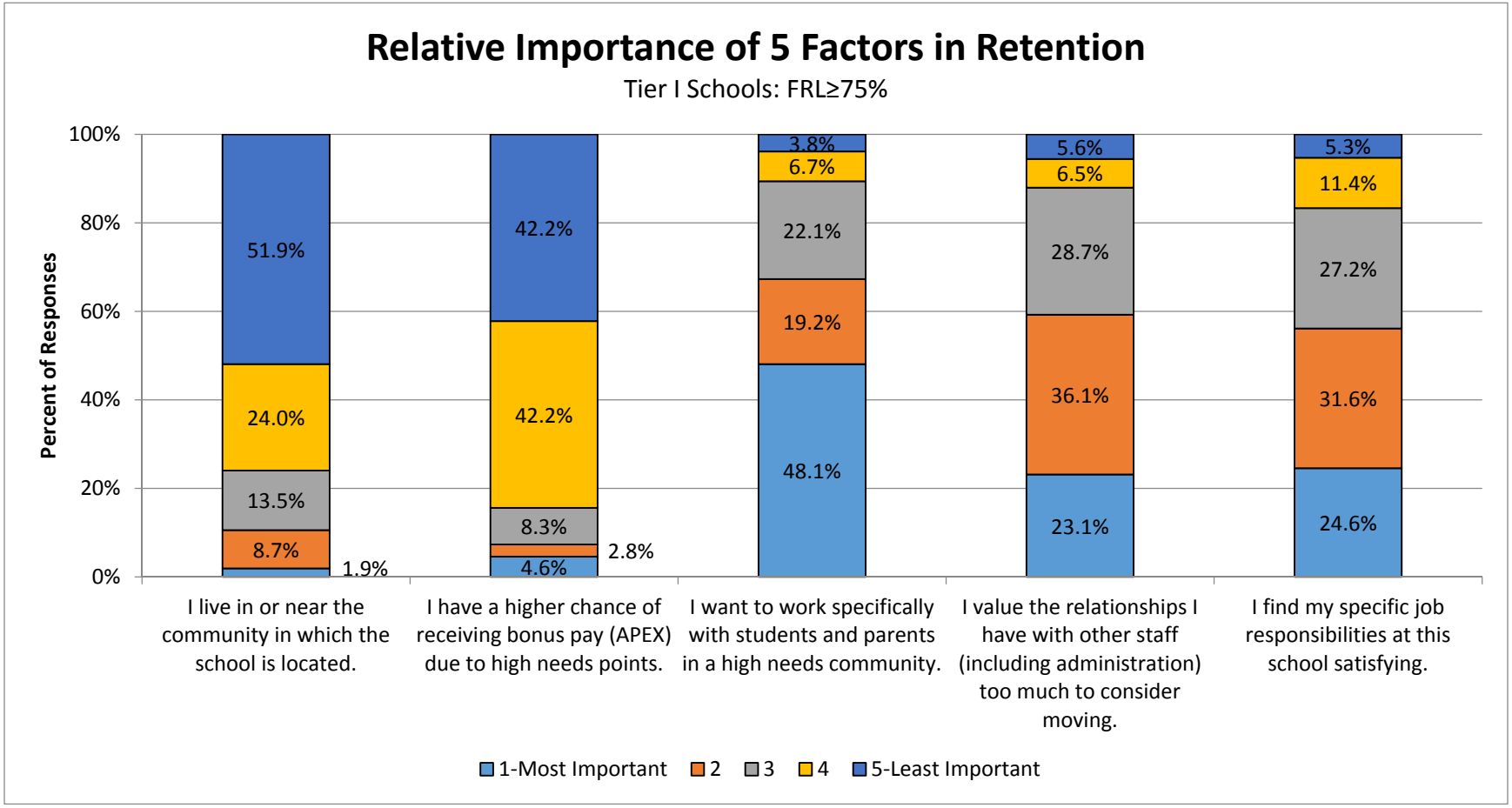


Figure 7.8: Reasons for Retention: Relative Rankings from Tier I Responses

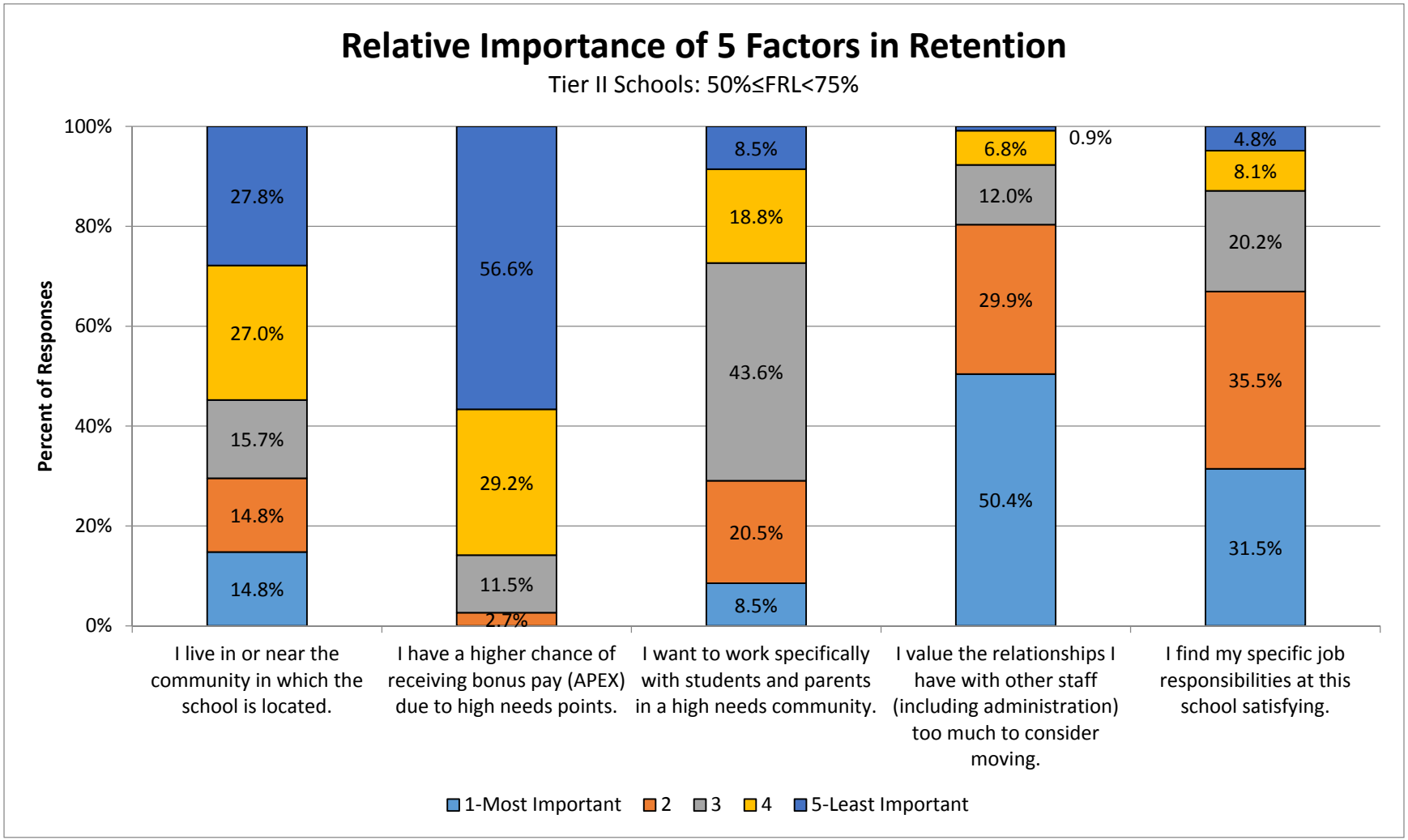


Figure 7.9: Reasons for Retention: Relative Rankings from Tier II Responses

The biggest consistency in the responses of both Tier I and Tier II respondents was that the APEX award was one of the least important factors in their continued employment in their current school. Respondents at Tier I schools were more likely to list APEX considerations as the most important factor than teachers in Tier II schools. However, only 7.4% of respondents placed APEX in the highest two categories of importance. The impact of APEX appears to have relatively little importance among teachers who choose to continue to serve in high-needs schools.

Conclusion and Considerations

The Knox County Schools strategic compensation program, Advance, Perform, Excel (APEX), was implemented to reward high quality instruction, student achievement, teacher leadership and continued service in high-needs schools. There is some evidence that some of these goals are being met, but the impact of APEX on these outcomes is difficult to quantify.

There is evidence that the quality of instruction of professional level certified staff has increased since APEX was implemented in SY1112. The mean difference between professional and apprentice level staff has increased each year since APEX was launched. There is also evidence that the number of teachers involved in leadership activities in their building increases as we move further from the initial implementation of APEX. Teacher retention data, however, is mixed. Multiple survey responses indicate that APEX is not a motivator for teacher retention. The retention data also seems to follow patterns. Retention of the most effective and least effective teachers decreased in the second year of APEX, but then increased in the third year. It is possible that trends in teacher retention are driven by larger external factors, such as the local job market.

Student outcome data is also mixed. Although there is a change in the growth trajectory of high schools after the implementation of APEX, the effect may be influenced more by the TEAM observation process than the strategic compensation program. It is important to note that despite any change in trajectories, high schools are still on a downward trajectory regarding TVAAS gains.

The key piece of evidence that strategic compensation may not be the driver for any increases is in the TNCRED survey responses. The most consistent and troubling aspects of the TNCRED data revolve around the perceptions on the impact of strategic compensation. The majority of respondents feel that the Knox County Schools strategic compensation initiative is having negative effects at their school. This trend is true across the three years for which data is available (SY1112, SY1213, and SY1314). A small percentage of respondents perceive that strategic compensation is helping teachers feel more satisfied with their jobs. Similar percentages of respondents feel that strategic compensation is helping teachers feel more valued as professionals. These were all key goals of the strategic compensation initiative. Purposeful changes would have to be made to APEX to change these perceptions. Any redesign of APEX must attempt to find the root cause of these responses and eliminate it.

The easiest recommendations to make regarding changes in APEX are based on longitudinal teacher responses collected through the TNCRED survey.

Table 7.6: Relative Importance of Metrics for Strategic Compensation

Rate how important each of the following factors should be in determining performance-based bonuses SY1314 Responses					
Question	**Not important**	**Low importance**	**Moderate Importance**	**Very Important**	**No Answer**
Teaching in hard to staff schools	7%	15%	40%	35%	3%
Teaching in hard to staff fields	8%	19%	40%	30%	3%
Helping other teachers improve their professional practice	3%	15%	49%	29%	3%
Working with students outside of class time	8%	26%	41%	22%	3%
The outcome of classroom observation completed under the Tennessee teacher evaluation system	7%	23%	50%	17%	3%
Final Overall rating under the Tennessee's teacher evaluation system	9%	25%	48%	15%	3%
High test scores on a standardized test	15%	33%	36%	13%	3%
Students' gains on TCAP as measured by TVAAS.	15%	33%	36%	13%	3%
National Board for Professional Teacher Standards certification	23%	38%	26%	9%	4%
Time spent in professional learning	15%	43%	34%	6%	3%

Pooling responses for moderate importance and very important places "helping teachers improve their instructional practice" as the most favorable metric. Arguably, this metric is already compensated through stipends for lead teachers in the TEAM schools and master/mentor teachers in the TAP schools. The metrics with the highest response rates for most important are service in hard to staff schools and hard to staff subject areas. KCS should consider making these two metrics key components in any revamped differentiated pay system.

Staffing Ratios

To what extent is class size associated with student performance? This paper will consider the distribution of class sizes through the scheduling data for the school year 2013-2014 (SY1314). From there we will investigate student growth as measured by the difference between each student's predicted scale score and actual scale score. This data exists for the eight courses for which there are state end-of-course (EOC) exams: Algebra I, Algebra II, English I, English II, English III, Biology I, Chemistry I, and United States History. We will not be looking at student performance data due to its inherent socio-economic bias. Other demographic features such as special education status and English Language Learner status would also be grounds for an achievement bias. The growth data eliminates these biases in that the predicted scores are produced by only considering the observed scores in relationship to each student's previous test history. This means that there is a theoretical level playing field when we consider growth in this manner. Of the 30,599 high school student schedules that could have resulted in an EOC test, 27,257 students took their EOC exam and had a sufficient test history to create a predictive score and subsequent growth score.

Methodology

To provide a way to effectively communicate the information with respect to class size, we have categorized each classroom by groups of five students. We will consider growth by the class size categories. This will be done overall as well as when crossed by class subjects.

Additionally, we will consider average student growth for the students of teachers who had at least one smaller class and at least one larger class to effectively control for the teacher.

Results: Distribution of Class Sizes

We were able to examine 1,479 scheduled classrooms. The most popular classroom size across the district for EOC courses in SY1314 was between 21 and 25 students. This can be seen in Figure 8.1.

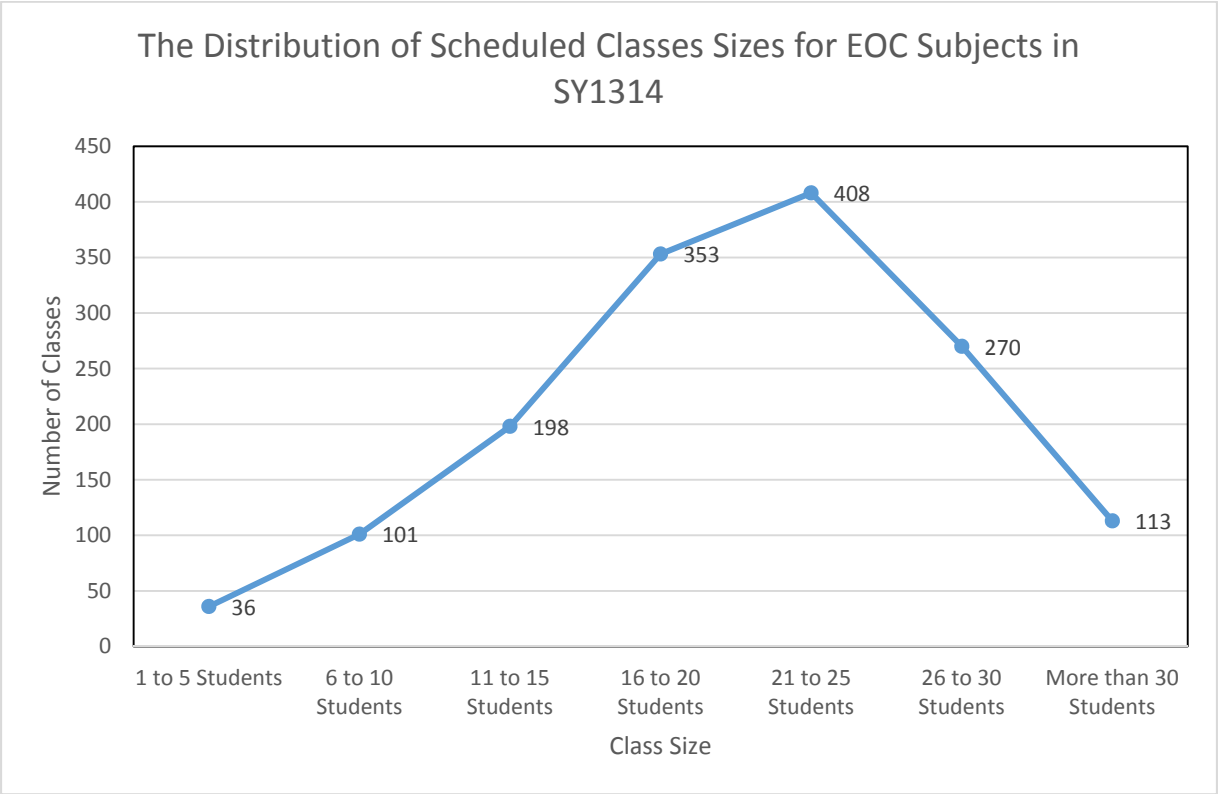


Figure 8.1: The Distribution of Class Sizes for EOC Subjects

A breakdown of these classes by size and subject can be found in Table 8.1 and a breakdown by size and school can be seen in Table 8.2. It should be noted that there are a variety of reasons for a class being the size that it is. Some, but not all, of the smaller classes were for special populations of students. A breakdown of the classes by size and subject and school is available in Appendix C.1. This appendix also includes the row percentages for each class size category. It is there that we can see that 27.6% of the classes had between 21 and 25 students and that three of the nine Algebra I classes with five or fewer students were at L & N STEM academy. This portion of the report is for background information as the creation of classes is not within the scope of this study.

Table 8.1: Number of Classes by Class Size and Subject

Subject	Number of Classes							
	1 to 5 Students	6 to 10 Students	11 to 15 Students	16 to 20 Students	21 to 25 Students	26 to 30 Students	More than 30 Students	Total Number of Classes
	Count	Count	Count	Count	Count	Count	Count	Count
Algebra I	9	21	42	68	54	17	8	219
Algebra II	1	9	30	38	54	40	8	180
Biology I	6	9	15	46	62	47	13	198
Chemistry I	1	6	31	61	48	29	17	193
English I	6	21	28	49	52	40	20	216
English II	2	12	20	43	61	35	21	194
English III	10	19	20	29	41	28	10	157
US History	1	4	12	19	36	34	16	122
Total	36	101	198	353	408	270	113	1479

Table 8.2: Number of Classes by Class Size and School

School	Number of Classes							
	1 to 5 Students	6 to 10 Students	11 to 15 Students	16 to 20 Students	21 to 25 Students	26 to 30 Students	More than 30 Students	Total Number of Classes
	Count	Count	Count	Count	Count	Count	Count	Count
Austin-East	5	21	33	18	3			80
Bearden	3	3	13	28	42	40	19	148
Carter		7	14	30	28	10	1	90
Central	1	4	8	31	40	13	3	100
Farragut	1	3	2	34	49	35	2	126
Fulton	9	11	27	31	20	1		99
Gibbs	1	6	5	15	29	23	8	87
Halls	2	8	13	36	34	17	2	112
Hardin Valley		11	21	12	22	32	44	142
Karns High	1	7	15	33	37	15		108
L & N STEM	3	2	3	4	10	12	9	43
Powell	1	4	14	25	31	29	15	119
South-Doyle	2	10	11	21	35	25	6	110
West	7	4	19	35	28	18	4	115
Total	36	101	198	353	408	270	113	1479

Results: Student Growth by Class Size

The mean growth of all 27,423 students over all of the schools and all of the subjects was just a little over 3.6 scale score points. The growth results varied quite a bit between schools and subjects. A general breakdown can be found in Table 8.3. We can note that Austin-East had the highest school mean growth with 9.8 scale score points while Chemistry I had the highest subject means growth with 9.6 scale score points.

Table 8.3: Growth by School and Subject

School	Subject								
	Algebra I	Algebra II	Biology I	Chemistry I	English I	English II	English III	US History	Total
	Growth	Growth	Growth	Growth	Growth	Growth	Growth	Growth	Growth
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Austin-East	33.2	29.5	6.3	-22.8	5.4	10.5	11.4	2.8	9.8
Bearden	-6.7	13.8	-.3	9.3	-1.9	-1.9	6.0	9.5	2.9
Carter	21.0	7.7	4.9	9.0	2.3	.0	5.3	9.0	7.5
Central	-14.0	-1.0	4.5	20.2	-1.9	1.0	4.7	-2.5	1.1
Farragut	4.3	16.9	3.8	14.2	.3	1.5	7.4	6.8	6.9
Fulton	-9.7	.6	-15.6	-.4	-2.8	2.0	8.9	6.4	-.7
Gibbs	-.3	4.7	-2.9	8.0	-4.4	.4	7.2	7.2	1.6
Halls	16.9	20.3	5.5	17.6	-2.5	2.3	4.2	7.7	8.4
Hardin Valley	-6.1	13.7	2.3	11.6	-1.0	-.2	4.2	4.1	3.4
Karns	-11.6	2.0	2.2	-1.5	-3.1	1.8	1.8	2.5	-.8
L & N STEM	-18.3	-7.2	-13.0	.2	-1.0	-6.0	5.9	5.4	-3.9
Powell	8.2	11.3	-1.0	15.1	-2.6	-1.4	3.1	6.6	4.6
South-Doyle	6.4	15.5	6.6	4.3	1.5	2.9	2.0	7.8	5.8
West	-6.3	-9.4	-8.9	22.4	-.9	5.2	4.5	7.5	.8
Total	.7	9.3	1.0	9.6	-1.2	.9	5.2	5.9	3.6

This information is expanded in Appendix C.2 to include student counts. It should be noted that these growth numbers may not exactly match the numbers on the TVAAS site. There are a few reasons why this is the case. One reason is that some students who are not actually scheduled for a course may take that course's exam. As an example, the TVAAS site notes 4,035 students took the exam in Biology I while Appendix C.2 notes that we had 3,873 students linked to schedules. Some of the missing students took the exam for Kelly Volunteer Academy where student scheduling is a little bit different from the rest of the schools and is therefore not

included in this study. It should also be noted that students who actually tested at Richard Yoakley or Byington Solway had their results attributed to one of the traditional high schools.

Having noted the breakdown of growth by school and subject, we will turn our attention to the focus of this study, the association of class size and growth. Table 8.4 and Figure 8.2 represent this data.

Table 8.4: Growth by Class Size

	Class Size							
	1 to 5 Students	6 to 10 Students	11 to 15 Students	16 to 20 Students	21 to 25 Students	26 to 30 Students	More than 30 Students	Total
	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean
All Subjects	-15.6	-5.9	3.1	3.7	4.4	4.0	3.9	3.6
Difference From Total	-19.2	-9.5	-0.6	0.1	0.7	0.3	0.3	0.0

The results show negative growth for the first two class size categories. This indicates that the observed mean scale score was less than the predicted mean scale score in these two categories. The opposite was the case for the remaining categories where the results vary between 3.1 and 4.4 scale score points. The green line in Figure 8.2 represents the threshold for zero growth while the overall mean growth of 3.6 scale score points is represented by the dotted line. Are the means for the individual class sizes significantly different than zero? Are they significantly different from the overall mean of 3.6?

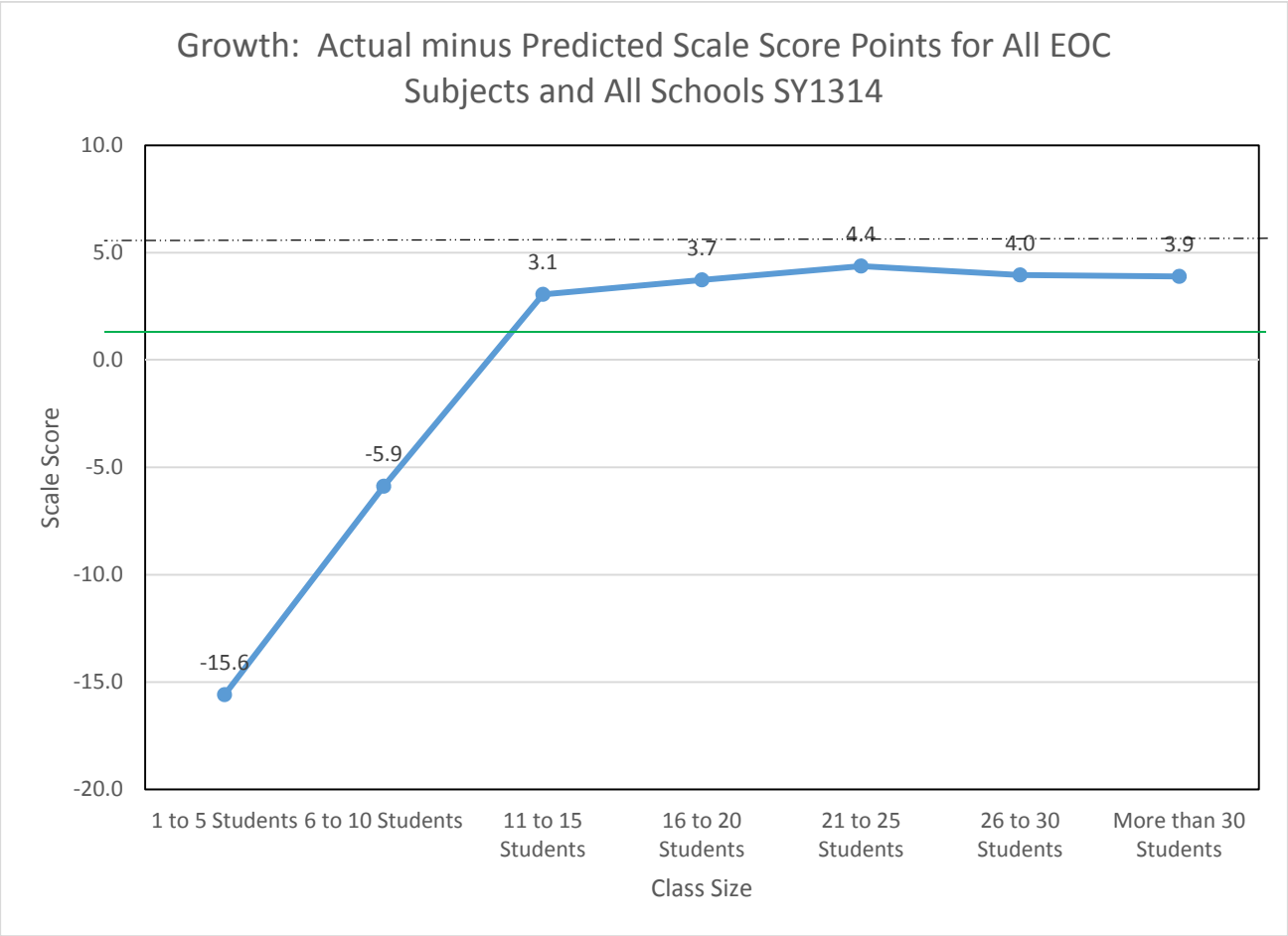


Figure 8.2: Growth by Class Size for All EOC Subjects

We performed t-tests on the individual class size categories to test for significance. Statistical significance is assigned for any p-value less than .05 as this indicates that the probability of a result this extreme happening by chance is less than one in 20. The criteria for the t-test results include the actual differences from the target, the number of data points, and the variation or spread of the data about the mean. In general, a larger number of data points require a smaller difference from the mean to be significant. Conversely, a small number of data points require a larger difference from the mean to achieve significance. The results for the hypothesis testing can be found in Table 8.5.

Table 8.5: t-test on Mean Class Size

Class Size	Count	Compared to 0		Compared to 3.6	
		t	p	t	p
1 to 5 Students	84	-2.439	0.017	-3.006	0.004
6 to 10 Students	742	-2.922	0.004	-4.722	0.000
11 to 15 Students	2327	3.191	0.001	-0.588	0.557
16 to 20 Students	5770	7.421	0.000	0.220	0.826
21 to 25 Students	8455	11.770	0.000	2.031	0.042
26 to 30 Students	6719	10.794	0.000	0.922	0.357
More than 30 Students	3326	9.194	0.000	0.644	0.520

The t-test results indicate that the class sizes of 1 to 5 students and 6 to 10 students performed significantly below zero (their predicted mean), as well as the overall mean growth of 3.6 scale score points. All of the other class sizes had means that were above zero in a statistically significant manner. While the class size of 21 to 25 students does not appear to be too much above the dotted line in Figure 8.2, this group performed above the overall mean in a statistically significant way. While there is not much difference between 4.4 and 3.62, significance was in a large part due to such a large number of data points (8,455). Finally, we note that the three smaller class size categories had growth means below the overall growth mean while the four larger class size categories had growth means above the overall growth mean.

Results: Student Growth by Class Size and Subject

With the initial dimension of class size examined, we will now consider the two dimensions of class size and subject to see if this further breakdown of the class size exhibits any trends. The crossing of class size and subject produced the mean gains found in Table 8.6.

Table 8.6: Mean Growth by Class Size and Subject

Subject	Class Size							
	1 to 5 Students	6 to 10 Students	11 to 15 Students	16 to 20 Students	21 to 25 Students	26 to 30 Students	More than 30 Students	Total
	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean	Growth Mean
Algebra I	-26.1	-.9	.5	1.8	.7	.5	-.4	.7
Algebra II	-89.7	.5	7.2	6.5	11.5	9.2	13.3	9.3
Biology I	-6.1	-14.8	-7.1	3.8	2.1	-.5	2.8	1.0
Chemistry I	50.6	8.2	7.1	8.1	9.6	13.3	8.8	9.6
English I	-18.2	-16.1	-1.1	-.5	.7	-1.2	-1.2	-1.2
English II	-35.9	-5.8	1.4	1.6	.3	1.0	1.8	.9
English III	-20.9	-6.6	9.4	4.4	6.5	5.4	6.5	5.2
US History	-3.4	6.5	5.0	5.8	6.3	6.4	4.7	5.9
Total	-15.6	-5.9	3.1	3.7	4.4	4.0	3.9	3.6

There are a few items that stand out in this table. Chemistry I exhibits positive growth for all class sizes. Running counter to the overall trend in classes with 1 to 5 students, the Chemistry I students in this category had the maximum gain on the chart. It turns out that there are only five students represented here out of the 84 students that are in this class size category. Appendix C.3 represents the counts in addition to the scale score growth means. Further investigation on our part found that these were five students from the same class at one of the high schools. The students in this class had a final class average of 84.6. Three of them ended up with a C for a grade while two of them ended up with a B. It appears that this one teacher did a phenomenal job in educating these five students. It also points out the possible dangers from attributing too much weight to this chart. The smallest mean was in the 1 to 5 students group in Algebra II. This came from two students, one of whom had the minimal possible scale score.

Aside from these outliers, we notice that the 21 to 25 students class size has positive growth for every subject. Even English I, the only subject with an overall negative mean growth, was positive in this category. We also note that for the two smallest class size categories that 13 out of the 16 cells exhibit negative growth.

It is possible to conduct a t-test on this data for each of the 56 cells. As was the case with class size alone, we tested against the mean of zero and the overall district mean of 3.6. The results of these tests can be found in Tables 8.7a and 8.7b.

Table 8.7a: t-test on Mean Class Size Crossed with Subjects against Mean Growth of Zero

Subject	Class Size													
	1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students	
	Growth		Growth		Growth		Growth		Growth		Growth		Growth	
	t	p	t	p	t	p	t	p	t	p	t	p	t	p
Algebra I	-2.146	0.050	-0.184	0.855	0.165	0.869	1.248	0.212	0.529	0.597	0.251	0.802	-0.149	0.882
Algebra II	-1.032	0.490	0.086	0.932	2.877	0.004	4.093	0.000	10.023	0.000	7.628	0.000	7.868	0.000
Biology I	-0.467	0.646	-1.727	0.089	-1.956	0.052	2.672	0.008	2.297	0.022	-0.546	0.585	2.639	0.009
Chemistry I	6.960	0.002	1.066	0.292	3.102	0.002	5.852	0.000	7.042	0.000	10.027	0.000	6.181	0.000
English I	-2.763	0.018	-4.270	0.000	-0.630	0.529	-0.636	0.525	1.033	0.302	-1.744	0.081	-1.795	0.073
English II	-0.811	0.503	-1.089	0.279	0.762	0.447	1.702	0.089	0.469	0.639	1.166	0.244	2.517	0.012
English III	-1.520	0.141	-1.388	0.167	4.191	0.000	2.477	0.014	6.158	0.000	5.712	0.000	5.089	0.000
US History	-0.175	0.890	1.762	0.088	2.091	0.038	4.043	0.000	7.478	0.000	8.241	0.000	4.897	0.000

Table 8.7b: t-test on Mean Class Size Crossed with Subjects against Mean Growth of 3.6

Subject	Class Size													
	1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students	
	Growth		Growth		Growth		Growth		Growth		Growth		Growth	
	t	p	t	p	t	p	t	p	t	p	t	p	t	p
Algebra I	-2.444	0.028	-0.912	0.363	-1.025	0.306	-1.258	0.209	-2.106	0.035	-1.520	0.129	-1.487	0.138
Algebra II	-1.074	0.477	-0.489	0.627	1.430	0.152	1.827	0.068	6.874	0.000	4.613	0.000	5.726	0.000
Biology I	-0.746	0.466	-2.150	0.035	-2.952	0.004	0.093	0.926	-1.606	0.109	-4.739	0.000	-0.809	0.419
Chemistry I	6.462	0.003	0.597	0.553	1.521	0.129	3.247	0.001	4.395	0.000	7.291	0.000	3.636	0.000
English I	-3.312	0.007	-5.229	0.000	-2.740	0.006	-4.952	0.000	-4.611	0.000	-6.797	0.000	-7.344	0.000
English II	-0.893	0.466	-1.763	0.082	-1.162	0.246	-2.042	0.042	-4.523	0.000	-3.180	0.002	-2.442	0.015
English III	-1.784	0.087	-2.155	0.033	2.581	0.010	0.436	0.663	2.735	0.006	1.912	0.056	2.256	0.025
US History	-0.362	0.779	0.770	0.433	0.567	0.572	1.529	0.127	3.177	0.002	3.594	0.000	1.104	0.270

When broken down by subject, the overall results somewhat match those of the class sizes by themselves from Table 8.5. Overall, there are more green cells representing significant positive growth to the right and more red cells representing significant negative growth to the left. We see that Chemistry exhibited statistically significant growth over many of the class sizes. Table 8.7b reveals the extent to which English I and English II performed below the district mean of 3.6 growth scale score points. As the tests are separate entities, it seemed reasonable to conduct an additional test and compare the individual subjects to their own means instead of zero or the group mean. The results of this test can be found in Table 8.8.

Table 8.8: t-test on Mean Class Size Crossed with Subjects against Subjects' Mean Growth

Subject	Mean Growth	Class Size													
		1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students	
		Growth		Growth		Growth		Growth		Growth		Growth		Growth	
		t	p	t	p	t	p	t	p	t	p	t	p	t	p
Algebra I	.74	-2.207	0.045	-0.333	0.740	-0.078	0.938	0.736	0.462	-0.009	0.992	-0.111	0.911	-0.423	0.673
Algebra II	9.29	-1.139	0.459	-1.390	0.169	-0.827	0.409	-1.723	0.085	1.942	0.052	-1.100	0.912	2.370	0.019
Biology I	.97	-0.542	0.594	-1.840	0.070	-2.222	0.027	1.981	0.048	1.251	0.211	-1.670	0.095	1.715	0.087
Chemistry I	9.59	5.641	0.005	-0.176	0.861	-1.088	0.277	-1.049	0.294	0.030	0.976	2.779	0.006	-0.560	0.576
English I	-1.17	-2.586	0.025	-3.960	0.000	0.052	0.958	0.759	0.448	2.856	0.004	-0.111	0.912	-0.002	0.998
English II	.89	-0.831	0.493	-1.255	0.213	0.289	0.773	0.782	0.435	-0.758	0.448	0.098	0.922	1.298	0.195
English III	5.15	-1.895	0.070	-2.480	0.014	1.900	0.059	-0.426	0.670	1.289	0.198	0.306	0.760	1.058	0.291
US History	5.91	-0.480	0.715	0.154	0.879	-0.398	0.691	-0.060	0.952	0.456	0.648	0.655	0.513	-1.296	0.196

Far fewer cells can be considered statistically significant when comparing each subject's class size growth mean to the subject's growth mean, but the pattern does not vary in that the areas with significant positive growth tend to be with class sizes of 16 or more while those with significant negative growth tend to be with the smaller class sizes. Cell counts and the differences from the subject means can be found in Appendix C.4.

Results: Student Growth by Class Size Controlling for Teachers

We have observed that smaller classes generally exhibit smaller or even negative growth when compared to larger classes. What about individual teachers who teach at least one smaller class and at least one larger class? Do these results remain when we control for the teachers? We chose to investigate this by examining the mean growth of students for teachers who had at least one class of 10 or fewer students as well as at least one class of at least 11 students in the same subject area. We were able to identify 49 teachers who had schedules that met our criteria. The results of comparing their student growth means by size can be found in Table 8.9.

Table 8.9: Mean Growth of Teachers with Smaller and Larger-sized Classes in Same Subject

Subject	Class Size of 10 or Fewer		Class size of 11 or More		Difference	t	p
	Growth		Growth		Growth		
	Mean	Count	Mean	Count	Mean		
Algebra I	-16.5	86	1.4	497	17.8	2.546	0.012
Algebra II	7.8	17	5.0	55	-2.8	-0.226	0.822
Biology I	-3.4	25	-7.9	286	-4.4	-0.555	0.579
Chemistry I	25.6	24	13.3	335	-12.2	-1.472	0.142
English I	-10.6	58	0.2	430	10.7	1.948	0.056
English II	3.0	26	3.1	134	0.1	0.01	0.992
English III	12.3	24	6.0	114	-6.3	-0.787	0.432
US History	6.2	17	0.6	50	-5.6	-0.788	0.434
Total	-3.2	277	2.3	1901	5.5	1.831	0.068

The results are mixed for this portion of the evaluation. Overall, students in the larger sized classes exhibited positive growth while the students in the smaller class sizes for these same teachers exhibited negative growth. Yet, the smaller classes performed better in five of the eight subjects. Algebra I was the main driver in the overall results. This subject had almost 27% of the students represented and the mean difference of 17.8 scale score points was not only the largest difference, but it was also statistically significant. As the results are not consistent across the subjects, we cannot conclude that the threshold of ten students affects individual teachers. Yet, there is a clear association with Algebra I for this threshold that is worthy of further inquiry.

Conclusions and Considerations

To what extent is class size associated with student performance? We used a growth measure for each student to remove any potential bias in the results. Overall, there is a significant association. Classes with ten or fewer students had a smaller growth mean than classes of 11 or more in a statistically significant way. In general, this was also the case when we broke the data down by subjects. When we considered individual teachers who had both smaller and larger size classes, the results were mixed, although Algebra I showed a statistically significant difference in favor of the larger classrooms.

It should be pointed out that association does not imply causation. While there is an association between class size and student growth, it should be pointed out that there are other considerations that go into class size. Some of the following considers may be affecting the growth results.

- Class sizes may be intentionally smaller at some locations for non-academic reasons. A reason could include limiting size for better classroom management.
- Schedule makers may intentionally or unintentionally overload the classrooms of the school's better teachers and under fill the classrooms of teachers whom they perceive to be less effective or not as student friendly.

Outside of purposeful student scheduling, there are other possible reasons for the relationship between class size and average growth.

- In small classrooms, students may feel more vulnerable because they may not have a peer that has a similar point of view or is struggling with a concept in the same way. If this is the case, students may be less inclined to actively participate in the classroom.
- Teachers may be more invigorated by the energy of a large class that a small class of students may not be able to provide. Our data suggests that this "critical mass" may be somewhere between 21 and 25 students in the classroom.
- When a few students are absent from a small class, a teacher may not want to push ahead for the sake of the missing students whose absence will be notable. In a larger class, the teacher may be more willing to press ahead with the curriculum and expect the missing students to catch up.
- When class sizes exceed 25 students, the mean growth starts to diminish. This may be due to the additional efforts required to maintain classroom management. On a related note, the larger the class, the more difficult it is for the teacher to establish a relationship with each student and his or her parents or guardians. This limits the amount of time available for student-led performance conferences, which have been shown to positively affect student performance.

Secondary Block Scheduling

In the early 1990's, Knox County Schools implemented block scheduling as the standard schedule structure in all of its high schools. For the implementation, all high schools converted from a traditional six-period school day to a 4x4 block. This change had major implications for students, teachers, and administrators. Students had to adjust to longer class periods and a faster pace because courses only met for half of the school year. Teachers had to adapt by incorporating new teaching strategies that are appropriate for a 90-minute block. In addition, teachers gained more planning and collaboration time under block scheduling. Administrators were able to offer more courses and were less constrained when creating their teacher and student schedules. This was due to the fact that teachers taught more classes over the course of the year even though they are actively teaching a smaller percentage of each day. It is for this reason that it generally takes more teachers to operate in a block scheduling scheme. The purpose of this report is to examine scheduling practices for the SY1314 and to determine any potential impact on staffing by moving back to a traditional schedule.

Methodology

In the past, an analysis of a traditional six-period schedule was conducted based on the staffing formula for each school. The staffing formula assumes that each staff member is to be fully scheduled and class sizes are fixed at a certain level, which is not always the case. This analysis differs from previous efforts because it incorporates actual class sizes and teacher utilization rates to account for current scheduling practices at each school. The SY1314 schedules at all high schools were analyzed to determine the potential impact on staffing by moving to a traditional six-period day. Since scheduling practices vary across the district, schedule structures were analyzed and converted to a standard time unit to facilitate the analysis. Additionally, some classes were removed in order to prevent them from skewing the results. These included classes that were not taught by KCS employees or classes that did not regularly meet during the school day. Once the data was prepared, a comparison analysis was conducted in three ways. In the first comparison, the average class size and teacher utilization rate calculated from SY1314 data were used to determine the number of staff members required under a traditional scheduling format and then compared to actual staffing levels. For the second comparison, the average class size was increased by 5% but the teacher utilization rate was held constant for each school in order to demonstrate the importance of class size. A third scenario was also considered where the average class size at each school was set to equal 25 students per class.

Results: Current State

In order to determine how teaching staff were being utilized across the district, several different measures were examined: SY1314 schedule structure, class size at each school, and the proportion of possible sections that were scheduled. All but one of the active high schools in SY1314, Dr. Paul L. Kelly Volunteer Academy, were still utilizing some form of block scheduling. Since block scheduling is not in use at Kelly Volunteer Academy, it was not included in this

analysis for SY1314. In all other high schools, some form of block scheduling is still in use. However, some schools have moved away from the 4x4 block scheduling structure (see Table 9.1). Four high schools are utilizing a combination of traditional and block scheduling (modified block) for a portion of their populations, and two other schools have implemented an alternating day block where classes meet every other day all year long. Since classes across the district are scheduled for different lengths of time, each course was converted to a standard unit of measurement. For the purposes of this analysis, the standard unit of measure is a block of 90 minutes daily for one semester.

Table 9.1: Schedule Structures Comparison in SY1314

School	4x4 Block	Modified Block	Alternating Day Block
Austin-East	✓		
Bearden	✓		
Carter	✓		
Central		✓	
South-Doyle		✓	
Farragut	✓		
Fulton		✓	
Gibbs	✓		
Halls		✓	
Hardin Valley	✓		
Karns	✓		
Powell	✓		
West			✓
L&N STEM			✓

Over 6,700 classes were included in the analysis and on average, class sizes varied from school to school. The mean class size in high schools across the district was 18.53 students per class with a standard deviation of 8.97, which indicates a high degree of variation in the individual class sizes across the district. The smallest average class size by school was 11.34 at Austin East High School and the largest was 21.27 at Bearden High School (see Figure 9.1). Tennessee state law provides guidelines for maximum class sizes in TCA § 49-1-104. The average number of students in non-vocational courses cannot exceed 30 students per class and no individual class can exceed 35 students. Similarly, vocational courses cannot exceed an average of 20 students per class and no individual class can be larger than 25 students. During SY1314, the average class size in Knox County for non-vocational classes was 19.45 and the average class size for vocational classes was 14.27 students per class. Both of these are well below the maximum class size limits outlined in state law. It should be noted that due to special scheduling circumstances at every school, there were a number of courses scheduled that had only one student enrolled in the class. In most cases, these courses were scheduled for administrative reasons and did not

meet on a regular basis. In order to prevent them from skewing the data set, those classes along with a few very large classes (such as band) were removed when calculating the school and district mean class sizes.

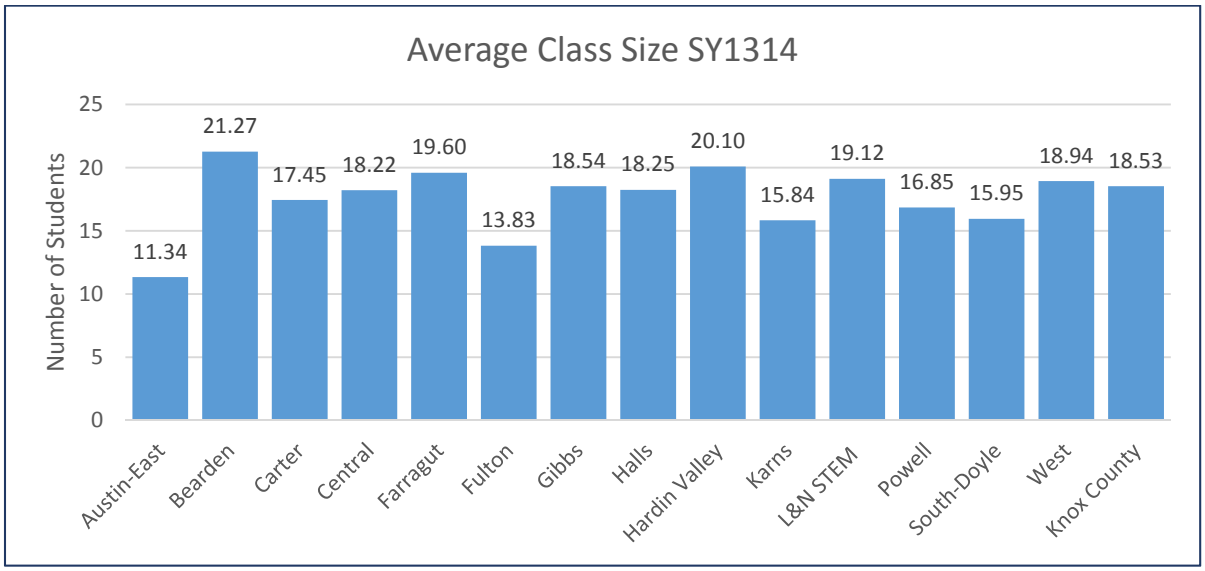


Figure 9.1: Average Class Sizes in SY1314

To get a sense of how core classes were scheduled in SY1314, the distribution of class sizes for Math and ELA was examined. A total of 1671 classes across the district were included and Table 9.2 shows the distributions for each school. Across the district, 52% of all Math and ELA classes contained 20 or fewer students but the distribution at some schools were quite different.

Table 9.2: Distribution of Math and ELA Classes by Number of Students

School	Number of Math and ELA Classes by Class Size, SY1314							
	0 to 5 students	6 to 10 students	11 to 15 students	16 to 20 students	21 to 25 students	26 to 30 students	31 to 35 students	36 or more students
Austin-East	5	31	52	17	4	0	0	0
Bearden	0	5	16	36	58	40	20	0
Carter	2	13	24	25	26	7	0	0
Central	3	5	11	36	37	11	2	1
Farragut	0	3	11	38	47	51	5	0
Fulton	4	13	33	43	22	5	0	0
Gibbs	0	8	10	16	27	23	9	0
Halls	0	5	11	40	37	20	2	0
Hardin Valley	4	12	16	22	32	36	43	0
Karns	4	8	15	41	34	13	0	0
L&N STEM	4	8	8	11	11	7	3	0
Powell	3	2	27	28	29	29	11	1
South-Doyle	6	4	15	23	33	21	3	0
West	0	9	27	48	25	22	3	0
Knox County	35	126	276	424	422	285	101	2

Results: Teaching Staff Course Load

For the purposes of this analysis, teaching staff was defined as any staff scheduled to teach a course in SY1314 with the exception of school counselors, school administrators, and librarians. Those three groups were omitted because the classes that they were scheduled to teach did not meet during the regular school day in most cases. In addition, a full time and fully scheduled teacher should be scheduled to a total of six semester block classes for a school year. With that in mind, the schedule capacity for each school was calculated by multiplying the total number of teaching staff by six. Staff members flagged part time were treated as half a position. There may be a few teachers who travel to more than two schools during a given segment of the school year, which could inflate the number of teachers at a given school but the effect would be small.

In order to get a sense of how teachers are currently being utilized by high schools across the district, the schedule capacity at each school was compared to the total number of scheduled semester blocks in SY1314. Figure 9.2 shows the rate at which each high school utilized the teachers at their location. The district as a whole utilized 93% of its full schedule capacity. Most schools were very close to or above the district rate with the exception of Austin-East that scheduled 83% of their capacity. There are a number of reasons why some teachers are not fully scheduled. These include unscheduled blocks of time that are sometimes allotted for serving as department head, mentor teacher, master teacher or general administrative tasks.

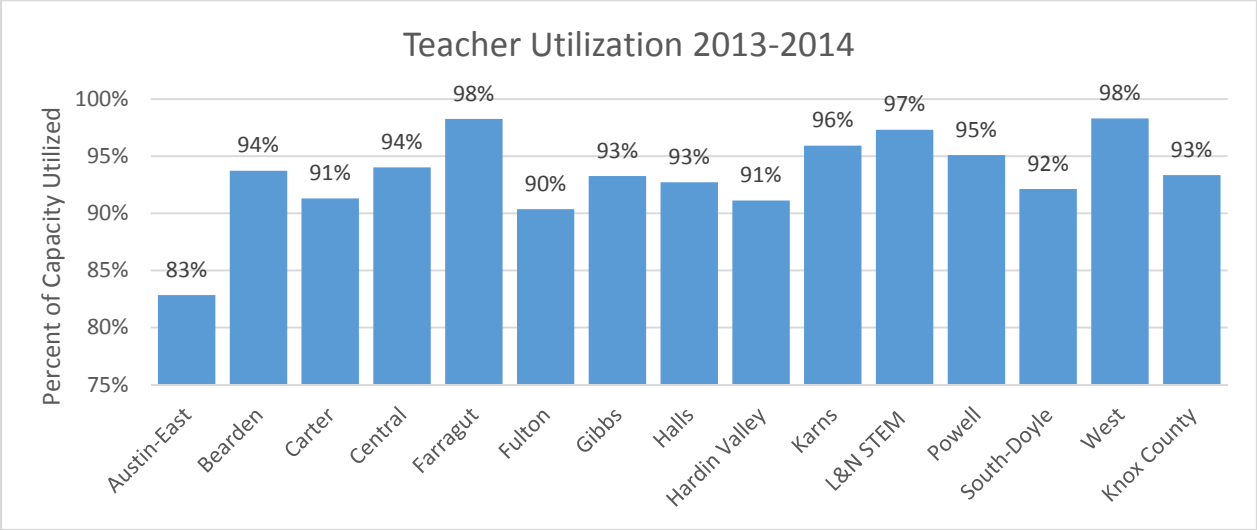


Figure 9.2: Teacher Utilization Rate in SY1314

Results: Block v. Six-Period Traditional Schedule

To project how many teachers would be required in a traditional six-period day, it was first determined how many classes would be needed at each school if a six-period day were implemented. As expected, the number of classes required for a six-period day at each school was 25% smaller than the number of classes required for block scheduling since students enroll in two fewer courses each year in a traditional schedule. After the total number of classes was

determined at each school, we then determined how many educators would be needed to teach that number of classes while taking the utilization rate into account for SY1314. The following equations were used to determine these values:

$$number\ of\ classes = \frac{school\ enrollment * 6}{avg.class\ size} \qquad Teachers\ needed = \frac{number\ of\ classes}{utilization\ rate * 5}$$

In the first projection, class sizes and teacher utilization remained at SY1314 levels. Using that model, we projected that it would take approximately 41 fewer educators across the district to accommodate a traditional six-period schedule. There is a projected decrease at every school with the exception of Fulton High School (see Table 9.2). Fulton and Austin-East High Schools both have very low average class sizes (14.34 and 11.53, respectively), which resulted in an increased number of required staff. It is important to note that this model is very sensitive to average class size and teacher utilization. That is, a small increase in class size or teacher utilization rate would result in a significant decrease in the number teaching staff required.

To understand how increased class size would impact staffing, two more models were considered. In the first of these scenarios, the class sizes at each school were increased by 5%. In that scenario, we projected that approximately 95 fewer staff members would be required across the district. The results from these calculations are listed in Table 9.3. In the second scenario, the average class size across the district was set to 25 students per class at every school. In that scenario, we projected that we would require 341 fewer teachers across the district. However, these results may be misleading because it would require unrealistic increases in class sizes, especially at Title 1 schools like Austin-East and Fulton. A percentage increase is more feasible because it takes the current average class size at each school into account.

Table 9.3: Comparison of Staffing Required for Block vs. Traditional Schedules

		SY1314 Class Size		SY1314 Class Size + 5%		SY1314 Class Size = 25	
School	Teaching Staff in SY1314	Projected Staff for 6 period day	Difference	Projected Staff for 6 period day	Difference	Projected Staff for 6 period day	Difference
Austin-East	68	67	-1	64	-4	64	-37
Bearden	117	112	-5	107	-10	107	-19
Carter	68	67	-1	64	-4	64	-20
Central	78	76	-2	72	-6	72	-20
South-Doyle	89	84	-5	80	-9	80	-29
Farragut	106	99	-7	94	-12	94	-24
Fulton	83	85	2	81	-2	81	-34
Gibbs	73	72	-1	69	-4	69	-18
Halls	87	85	-2	81	-6	81	-23
Hardin Valley	122	116	-6	111	-11	111	-23
Karns	92	90	-2	86	-6	86	-32
Powell	102	97	-5	92	-10	92	-33
West	79	76	-3	72	-7	72	-19
L&N STEM	31	29	-2	28	-3	28	-8
	Total		**-41**		**-95**		**-341**

Conclusions and Considerations

Average class sizes in Knox County high schools appear to be small when compared to state guidelines and the issue is more pronounced at some locations. Small class sizes across the district suggest educators are not being used to maximum effect. Therefore, the district may wish to investigate increasing class size regardless of whether a change is made to the schedule structure. If, however, a change in schedule structure from block scheduling to a traditional schedule were made, it would reduce the number of teaching staff required at almost every high school across the district resulting in significant cost reductions. Additionally, a small increase in class size coupled with a traditional schedule could result in even greater cost reductions. While the potential staffing cost savings is compelling, a change in scheduling is a fundamental change in the way schools are structured and it will have far reaching implications to almost all stakeholders. Some of the possible consequences outside of staffing include:

- Students in a traditional schedule are enrolled in fewer courses when compared to a block schedule which is one of the reasons why that schedule structure requires fewer staff. Since students would not be able to take as many classes over the course of their academic career, students would have less choice and flexibility when choosing their plan

of study. It is also likely that schools would be forced to limit the variety of courses that they offer.

- Since most courses in a traditional schedule span the entire year, schools would have less flexibility when building their master schedules.
- Teachers would be teaching a larger percentage of the school day and would have less time to engage in collaboration with their colleagues, planning, or participate in other school duties.
- Since all students in the district would be enrolled in their core academic subjects all year, it is likely that more textbooks would be required to operate a traditional schedule. Costs associated with buying additional textbooks would offset some of the cost savings.

Appendix A.1: Early Literacy

			At Least One On Target				Both Above Target			
			Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth	
			Mean	Mean	Mean	Count	Mean	Mean	Mean	Count
Grade One	Literacy Intervention Student	No	537.1	529.1	-8.0	592	593.6	595.3	1.8	2702
		Yes	538.5	528.1	-10.4	260	559.9	553.5	-6.4	142
	Difference (Yes Minus No)		1.4	-1.0	-2.4	p = .290	-33.7	-41.8	-8.1	p = .001
Grade Two	Literacy Intervention Student	No	573.6	569.7	-4.0	468	629.6	630.5	.9	2706
		Yes	576.3	573.2	-3.1	308	604.3	604.4	.0	161
	Difference (Yes Minus No)		2.7	3.6	0.9	p = .611	-25.2	-26.1	-0.9	p = .628
Grade Three	Literacy Intervention Student	No	720.9	717.7	-3.2	484	768.6	771.0	2.4	2459
		Yes	728.4	719.5	-8.9	396	744.6	742.4	-2.2	200
	Difference (Yes Minus No)		7.5	1.9	-5.7	p = .000	-24.1	-28.6	-4.6	p = .002
Grade Four	Literacy Intervention Student	No	35.2	32.3	-2.9	578	67.2	66.5	-.7	2532
		Yes	36.9	34.1	-2.8	350	56.7	54.4	-2.3	136
	Difference (Yes Minus No)		1.7	1.8	0.1	p = .874	-10.5	-12.1	-1.6	p = .157
Grade Five	Literacy Intervention Student	No	33.6	31.6	-2.0	565	69.5	64.8	-4.6	2574
		Yes	34.2	31.3	-2.9	311	54.9	51.0	-3.9	110
	Difference (Yes Minus No)		0.6	-0.3	-0.9	p = .349	-14.6	-13.9	0.7	p = .542

			At Least One On Target				Both Above Target			
			Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth	
			Mean	Mean	Mean	Count	Mean	Mean	Mean	Count
Grade One	Literacy Intervention Student	No	517.1	498.0	-19.1	164	569.1	563.6	-5.5	188
		Yes	511.0	485.2	-25.9	64	547.0	523.0	-24.0	32
	Difference (Yes Minus No)		-6.1	-12.9	-6.8	p = .083	-22.1	-40.6	-18.4	p = .001
Grade Two	Literacy Intervention Student	No	550.0	540.7	-9.3	277	614.4	610.5	-3.9	182
		Yes	543.2	529.1	-14.1	64	590.0	572.6	-17.5	18
	Difference (Yes Minus No)		-6.8	-11.6	-4.8	p = .107	-24.4	-38.0	-13.6	p = .036
Grade Three	Literacy Intervention Student	No	698.7	698.3	-.5	274	754.5	751.8	-2.6	162
		Yes	700.1	693.0	-7.2	52	736.0	723.1	-12.8	49
	Difference (Yes Minus No)		1.4	-5.3	-6.7	p = .142	-18.5	-28.7	-10.2	p = .020
Grade Four	Literacy Intervention Student	No	22.1	22.7	.6	188	48.5	45.6	-2.9	223
		Yes	14.9	16.6	1.8	56	37.3	34.4	-2.8	32
	Difference (Yes Minus No)		-7.3	-6.0	1.2	p = .536	-11.2	-11.2	0.0	p = 0.994
Grade Five	Literacy Intervention Student	No	20.2	18.3	-1.9	184	43.8	41.0	-2.8	235
		Yes	13.3	14.4	1.1	40	40.3	32.7	-7.6	54
	Difference (Yes Minus No)		-6.9	-3.9	3.0	p = .154	-3.5	-8.3	-4.8	p = .045

Appendix A.2: Early Literacy

Grade One	Literacy Intervention Student								Growth Difference for n > 4
	No				Yes				
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	mean
A. L. Lotts Elementary	592.9	593.7	.9	91	542.5	541.4	-1.2	14	-2.0
Adrian Burnett Elementary	556.5	568.1	11.6	79	527.1	513.0	-14.1	2	
Amherst Elementary	583.4	573.9	-9.5	127	536.2	506.3	-29.9	3	
Ball Camp Elementary	598.8	594.5	-4.3	64	546.5	525.2	-21.3	22	-17.0
Bearden Elementary	591.8	588.5	-3.3	52	546.9	555.8	8.9	9	12.1
Beaumont Elementary	587.7	584.1	-3.7	49	544.2	534.1	-10.1	31	-6.5
Belle Morris Elementary	562.9	559.6	-3.3	68					
Blue Grass Elementary	597.7	606.6	8.9	77	543.6	539.4	-4.2	13	-13.1
Bonny Kate Elementary	586.3	569.7	-16.6	43	547.0	526.2	-20.8	5	-4.1
Brickey-McCloud Elementary	593.7	593.0	-.7	135	545.6	525.9	-19.7	7	-19.0
Carter Elementary	576.1	567.5	-8.6	71					
Cedar Bluff Elementary	577.5	579.0	1.4	128	546.0	531.0	-14.9	38	-16.4
Christenberry Elementary	561.9	559.5	-2.4	81					
Copper Ridge Elementary	582.3	574.6	-7.8	78					
Corryton Elementary	563.0	556.2	-6.8	29					
Dogwood Elementary	574.4	576.5	2.1	69	533.6	521.8	-11.8	40	-13.9
East Knox County Elementary	563.2	553.2	-10.0	73	533.7	533.4	-.3	5	9.8
Farragut Primary	594.1	592.9	-1.2	254					
Fountain City Elementary	599.6	600.8	1.2	26	551.2	552.5	1.2	23	0.0
Gap Creek Elementary	578.6	588.0	9.4	21					
Gibbs Elementary	577.4	578.4	.9	94	532.4	525.2	-7.2	31	-8.1
Green Elementary	567.7	563.7	-4.0	36	541.4	517.4	-24.0	14	-20.0
Halls Elementary	572.3	580.1	7.8	104	538.2	540.9	2.8	16	-5.0
Hardin Valley Elementary	592.1	583.7	-8.3	131					
Inskip Elementary	572.4	568.3	-4.2	80	560.7	558.4	-2.3	5	1.8
Karns Elementary	570.5	572.5	2.0	181	518.7	465.0	-53.7	1	
Lonsdale Elementary	552.8	538.6	-14.1	64					
Maynard Elementary	536.6	543.0	6.4	26	512.5	477.0	-35.5	1	
Mooreland Heights Elementary	541.8	540.3	-1.5	54	529.2	486.0	-43.2	2	
Mount Olive Elementary	590.7	581.3	-9.4	44					
New Hopewell Elementary	581.6	584.4	2.9	28	535.9	514.9	-21.1	15	-24.0
Northshore Elementary	610.8	614.8	4.0	111	551.3	545.2	-6.1	26	-10.2
Norwood Elementary	566.7	568.8	2.1	61	536.9	520.1	-16.8	32	-18.9
Pleasant Ridge Elementary	578.3	571.9	-6.4	83					
Pond Gap Elementary	595.3	588.1	-7.1	14	550.5	528.6	-21.9	30	-14.7
Powell Elementary	575.7	580.6	4.8	139	558.1	553.8	-4.2	13	-9.1
Ritta Elementary	569.4	561.1	-8.3	84					
Rocky Hill Elementary	603.5	599.2	-4.3	102					
Sarah Moore Greene	555.4	552.3	-3.1	108					
Sequoyah Elementary	603.5	611.4	7.9	59	558.0	558.5	.5	8	-7.4
Shannondale Elementary	596.5	597.9	1.5	57					
South Knoxville Elementary	576.6	572.5	-4.1	12	533.2	510.4	-22.8	10	-18.7
Spring Hill Elementary	569.6	564.5	-5.2	49	541.0	530.7	-10.3	28	-5.1
Sterchi Elementary	585.9	584.0	-1.9	56					
Sunnyview Primary	570.9	571.5	.6	75	538.0	516.1	-21.9	7	-22.5
West Haven Elementary	589.9	589.7	-.2	39	534.0	522.2	-11.9	20	-11.7
West Hills Elementary	576.4	589.9	13.5	102	529.6	508.9	-20.7	27	-34.2
West View Elementary	550.9	548.1	-2.8	38					
Knox County	**579.7**	**578.6**	**-1.1**	**3646**	**541.6**	**529.5**	**-12.1**	**498**	**-11.0**

Grade Two	Literacy Intervention Student								Growth Difference for n > 4
	No				Yes				
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	mean
A. L. Lotts Elementary	634.5	633.2	-1.3	111	561.4	555.7	-5.7	11	-4.4
Adrian Burnett Elementary	599.2	592.6	-6.7	85	572.0	550.0	-22.0	12	-15.3
Amherst Elementary	618.6	621.3	2.8	102	585.4	579.6	-5.9	18	-8.6
Ball Camp Elementary	622.7	629.6	6.8	80	580.1	585.5	5.4	19	-1.4
Bearden Elementary	620.1	619.0	-1.2	56					
Beaumont Elementary	629.9	624.1	-5.8	60	578.3	565.0	-13.3	24	-7.5
Belle Morris Elementary	604.9	581.8	-23.1	8	593.2	591.2	-1.9	80	21.2
Blue Grass Elementary	634.2	631.7	-2.5	93	594.4	589.2	-5.2	13	-2.7
Bonny Kate Elementary	606.8	608.7	1.9	63	577.0	564.5	-12.5	4	
Brickey-McCloud Elementary	629.5	626.7	-2.7	116	584.9	576.1	-8.8	25	-6.1
Carter Elementary	605.3	607.1	1.8	67	597.0	566.0	-31.0	1	
Cedar Bluff Elementary	609.4	607.7	-1.7	120	589.4	580.3	-9.1	19	-7.4
Christenberry Elementary	589.8	593.5	3.7	72					
Copper Ridge Elementary	601.3	599.9	-1.4	79					
Corryton Elementary	616.4	619.0	2.6	36					
Dogwood Elementary	606.6	602.3	-4.3	62	571.9	564.9	-7.0	36	-2.8
East Knox County Elementary	595.5	594.2	-1.2	55	566.7	553.8	-12.9	6	-11.7
Farragut Primary	627.5	625.0	-2.5	327	584.5	574.8	-9.7	5	-7.2
Fountain City Elementary	615.2	609.0	-6.2	29	592.7	596.7	4.0	21	10.2
Gap Creek Elementary	605.9	612.5	6.6	14					
Gibbs Elementary	609.6	621.0	11.4	92	577.7	593.0	15.3	29	3.9
Green Elementary	589.8	589.1	-.7	31	568.3	557.6	-10.7	20	-10.0
Halls Elementary	612.6	608.6	-4.0	101	558.4	540.6	-17.8	5	-13.9
Hardin Valley Elementary	625.9	623.7	-2.2	163					
Inskip Elementary	613.7	611.5	-2.2	68	556.0	549.3	-6.8	4	
Karns Elementary	606.3	609.4	3.1	165					
Lonsdale Elementary	583.0	574.7	-8.3	61	586.9	573.0	-13.9	10	-5.5
Maynard Elementary	600.5	587.9	-12.6	20					
Mooreland Heights Elementary	587.4	591.1	3.8	65	584.8	562.0	-22.8	1	
Mount Olive Elementary	603.9	603.2	-.7	40					
New Hopewell Elementary	615.4	613.8	-1.6	33	572.0	549.3	-22.7	10	-21.0
Northshore Elementary	633.8	630.0	-3.8	132	588.7	571.8	-16.9	9	-13.1
Norwood Elementary	607.7	611.5	3.8	63	565.4	568.4	3.0	21	-0.9
Pleasant Ridge Elementary	602.0	593.8	-8.2	46					
Pond Gap Elementary	629.6	624.6	-4.9	10	580.1	586.2	6.1	28	11.0
Powell Elementary	613.1	612.8	-.3	114	583.0	584.0	1.0	22	1.3
Ritta Elementary	609.8	607.2	-2.6	107	571.3	579.2	7.9	6	10.5
Rocky Hill Elementary	633.1	638.6	5.5	114					
Sarah Moore Greene	584.6	579.9	-4.6	93	563.9	558.5	-5.4	2	
Sequoyah Elementary	649.2	655.1	5.9	74					
Shannondale Elementary	624.2	626.8	2.6	67	566.8	532.0	-34.8	1	
South Knoxville Elementary	607.1	604.7	-2.4	6	583.0	571.5	-11.5	8	-9.1
Spring Hill Elementary	611.2	603.8	-7.4	42	585.2	580.4	-4.8	30	2.6
Sterchi Elementary	617.0	624.8	7.8	37	592.3	598.5	6.2	6	-1.6
Sunnyview Primary	603.3	605.1	1.8	86	588.9	620.0	31.1	1	
West Haven Elementary	630.4	624.5	-5.9	35	571.1	557.4	-13.7	17	-7.8
West Hills Elementary	630.9	629.7	-1.1	100	582.2	583.4	1.2	25	2.4
West View Elementary	600.0	590.3	-9.6	33	573.3	584.0	10.8	2	
Knox County	**615.5**	**614.8**	**-.7**	**3633**	**581.1**	**577.2**	**-3.9**	**551**	**-3.2**

Grade Three	Literacy Intervention Student								Growth Difference for n > 4
	No				Yes				
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	mean
A. L. Lotts Elementary	763.5	769.6	6.1	108	725.6	717.3	-8.3	9	-14.3
Adrian Burnett Elementary	737.8	737.2	-.5	91	718.6	709.9	-8.6	16	-8.1
Amherst Elementary	757.7	757.1	-.5	91	726.9	713.3	-13.6	21	-13.1
Ball Camp Elementary	765.6	771.3	5.8	60	731.8	730.6	-1.2	28	-6.9
Bearden Elementary	761.0	754.4	-6.6	46					
Beaumont Elementary	765.3	766.3	1.0	56	712.1	703.9	-8.2	35	-9.2
Belle Morris Elementary	732.5	745.0	12.5	1	744.4	744.3	-.1	62	
Blue Grass Elementary	769.2	772.6	3.4	96	729.5	708.8	-20.7	13	-24.1
Bonny Kate Elementary	756.3	759.2	2.9	48	728.7	724.0	-4.7	8	-7.5
Brickey-McCloud Elementary	766.1	774.0	7.8	112	736.2	726.1	-10.2	48	-18.0
Carter Elementary	752.6	755.2	2.6	82					
Cedar Bluff Elementary	751.5	756.1	4.6	151	729.4	724.4	-5.0	34	-9.6
Chilhowee Intermediate	750.5	759.9	9.5	44	733.3	728.3	-5.1	35	-14.5
Christenberry Elementary	734.9	744.3	9.4	63	730.8	736.8	6.1	17	-3.3
Copper Ridge Elementary	749.0	739.7	-9.4	94	728.1	702.0	-26.1	1	
Corryton Elementary	756.9	750.6	-6.3	21	740.6	725.5	-15.1	6	-8.8
Dogwood Elementary	739.1	734.0	-5.1	74	729.1	726.3	-2.7	39	2.4
East Knox County Elementary	736.6	725.5	-11.1	73					
Farragut Intermediate	771.7	777.4	5.6	273	735.4	725.7	-9.7	28	-15.3
Fountain City Elementary	749.6	745.6	-4.0	35	733.4	722.0	-11.4	15	-7.4
Gap Creek Elementary	757.6	787.4	29.8	11					
Gibbs Elementary	756.1	758.4	2.3	87	735.5	726.2	-9.3	30	-11.6
Green Elementary	737.4	722.4	-15.0	20	721.7	703.3	-18.4	22	-3.4
Halls Elementary	754.7	752.1	-2.6	102	728.3	718.0	-10.3	2	
Hardin Valley Elementary	761.9	757.5	-4.3	137	723.4	728.0	4.6	2	
Inskip Elementary	739.0	741.6	2.6	65	699.2	680.5	-18.7	4	
Karns Elementary	756.6	757.7	1.1	181					
Lonsdale Elementary	730.0	729.3	-.6	46					
Maynard Elementary	736.4	752.6	16.2	38					
Mooreland Heights Elementary	752.3	750.5	-1.8	49	732.9	736.7	3.7	3	
Mount Olive Elementary	740.9	740.5	-.3	36					
New Hopewell Elementary	747.8	746.8	-1.0	24	729.8	717.7	-12.1	7	-11.1
Northshore Elementary	774.5	780.4	5.9	99	737.9	732.6	-5.3	27	-11.2
Norwood Elementary	739.7	738.6	-1.1	77	730.1	721.8	-8.2	19	-7.2
Pleasant Ridge Elementary	755.8	747.4	-8.4	41	739.3	745.0	5.7	1	
Pond Gap Elementary	756.7	768.2	11.6	9	735.0	734.3	-.7	46	-12.3
Powell Elementary	755.5	754.6	-.9	123	731.8	719.6	-12.2	28	-11.4
Ritta Elementary	749.5	753.7	4.2	81	714.5	712.0	-2.5	1	
Rocky Hill Elementary	776.8	779.1	2.3	89	729.2	718.3	-10.9	6	-13.2
Sarah Moore Greene	721.0	712.4	-8.6	85	704.8	682.0	-22.8	1	
Sequoyah Elementary	776.9	777.8	.9	66					
Shannondale Elementary	760.0	768.4	8.4	67					
South Knoxville Elementary	762.7	768.1	5.5	14	718.7	711.8	-6.8	6	-12.3
Spring Hill Elementary	762.1	759.8	-2.3	12	739.3	729.5	-9.8	34	-7.6
Sterchi Elementary	753.9	760.1	6.2	44	735.7	717.1	-18.5	7	-24.8
West Haven Elementary	748.8	751.1	2.3	34	717.2	717.8	.6	19	-1.7
West Hills Elementary	763.9	758.3	-5.6	88	732.2	710.6	-21.6	16	-16.0
West View Elementary	738.6	735.1	-3.5	35	727.0	737.0	10.0	1	
Knox County	**755.4**	**756.5**	**1.1**	**3379**	**731.5**	**724.4**	**-7.1**	**697**	**-8.2**

Grade Four	Literacy Intervention Student								Growth Difference for n > 4
	No				Yes				
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	mean
A. L. Lotts Elementary	65.4	67.4	2.0	115	35.1	37.6	2.5	10	0.5
Adrian Burnett Elementary	50.1	49.7	-.5	67	36.5	33.2	-3.3	15	-2.8
Amherst Elementary	64.1	57.6	-6.5	88	43.4	40.4	-3.0	21	3.5
Ball Camp Elementary	65.2	66.6	1.4	63	43.2	45.6	2.4	13	1.0
Bearden Elementary	56.1	61.0	5.0	44	26.0	24.0	-2.0	1	
Beaumont Elementary	57.5	60.6	3.1	47	28.6	28.3	-.3	25	-3.4
Belle Morris Elementary	67.8	55.2	-12.6	5	43.6	43.3	-.3	63	12.3
Blue Grass Elementary	66.5	65.1	-1.4	85	44.8	40.7	-4.2	12	-2.8
Bonny Kate Elementary	57.1	57.0	-.1	41	50.7	52.2	1.5	6	1.6
Brickey-McCloud Elementary	63.8	62.5	-1.3	113	39.8	31.2	-8.6	21	-7.3
Carter Elementary	56.0	54.1	-1.9	88					
Cedar Bluff Elementary	61.2	59.3	-1.9	167	34.9	30.6	-4.3	15	-2.3
Chilhowee Intermediate	56.6	56.9	.4	59	39.2	37.6	-1.6	16	-1.9
Christenberry Elementary	55.3	54.8	-.5	57	51.8	35.5	-16.3	4	
Copper Ridge Elementary	56.4	59.4	3.0	77					
Corryton Elementary	63.2	57.9	-5.4	31					
Dogwood Elementary	48.9	42.9	-6.0	65	33.3	30.2	-3.1	36	2.9
East Knox County Elementary	43.0	38.4	-4.6	68					
Farragut Intermediate	69.3	67.9	-1.4	246	51.7	50.6	-1.1	47	0.3
Fountain City Elementary	58.9	53.8	-5.1	51	41.6	34.7	-7.0	22	-1.8
Gap Creek Elementary	58.6	58.8	.2	9					
Gibbs Elementary	60.4	60.3	-.2	101	37.9	39.3	1.4	22	1.5
Green Elementary	38.2	35.3	-2.9	36	17.1	21.1	4.0	8	6.9
Halls Elementary	53.0	50.8	-2.2	151	33.5	27.0	-6.5	2	
Hardin Valley Elementary	64.6	61.3	-3.3	149	41.0	37.0	-4.0	1	
Inskip Elementary	50.7	48.1	-2.7	60	22.3	18.5	-3.8	8	-1.1
Karns Elementary	57.5	59.9	2.4	207	49.5	46.5	-3.0	2	
Lonsdale Elementary	38.5	33.0	-5.6	38					
Maynard Elementary	38.7	30.4	-8.3	25	22.0	24.5	2.5	2	
Mooreland Heights Elementary	49.9	40.9	-9.0	48	34.7	15.0	-19.7	3	
Mount Olive Elementary	58.7	53.1	-5.6	34					
New Hopewell Elementary	58.7	55.5	-3.3	36	36.3	37.4	1.1	8	4.4
Northshore Elementary	70.4	70.9	.5	107	46.5	46.9	.4	18	-0.1
Norwood Elementary	46.7	46.3	-.4	60	32.5	36.8	4.3	22	4.7
Pleasant Ridge Elementary	53.3	57.4	4.1	60	8.0	17.0	9.0	1	
Pond Gap Elementary	62.0	65.8	3.8	9	34.7	30.8	-3.9	37	-7.7
Powell Elementary	56.5	54.8	-1.6	100	38.0	33.8	-4.2	26	-2.6
Ritta Elementary	50.7	49.9	-.8	103	19.0	1.0	-18.0	1	
Rocky Hill Elementary	67.5	67.1	-.4	109					
Sarah Moore Greene	36.2	30.9	-5.3	109	27.5	18.0	-9.5	2	
Sequoyah Elementary	70.3	76.9	6.6	60	35.5	39.0	3.5	2	
Shannondale Elementary	62.7	61.5	-1.2	68					
South Knoxville Elementary	70.6	70.0	-.6	10	38.6	31.1	-7.5	10	-6.9
Spring Hill Elementary	53.1	52.5	-.6	26	42.6	39.6	-3.0	36	-2.4
Sterchi Elementary	63.7	65.2	1.5	56	49.3	51.7	2.3	6	0.9
West Haven Elementary	63.3	61.9	-1.3	39	32.0	31.2	-.8	12	0.5
West Hills Elementary	58.7	58.0	-.7	108	45.3	39.9	-5.4	18	-4.7
West View Elementary	44.9	51.9	7.0	26					
Knox County	**58.3**	**57.2**	**-1.1**	**3521**	**39.4**	**37.2**	**-2.2**	**574**	**-1.1**

Grade Five	Literacy Intervention Student								Growth Difference for n > 4
	No				Yes				
	Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		
	Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	mean
A. L. Lotts Elementary	63.9	62.0	-1.9	130	40.5	30.0	-10.5	4	
Adrian Burnett Elementary	46.8	42.0	-4.8	93	28.8	30.4	1.6	10	6.4
Amherst Elementary	60.0	53.0	-7.0	109	33.0	29.2	-3.8	19	3.2
Ball Camp Elementary	61.6	57.5	-4.1	65	32.9	28.0	-4.9	12	-0.8
Bearden Elementary	62.9	59.6	-3.3	51	16.0	1.0	-15.0	1	
Beaumont Elementary	64.9	57.6	-7.2	41	32.6	24.3	-8.2	23	-1.0
Belle Morris Elementary	30.1	25.8	-4.4	8	48.3	46.7	-1.6	56	2.8
Blue Grass Elementary	72.7	69.9	-2.8	129	38.6	35.5	-3.1	13	-0.3
Bonny Kate Elementary	58.3	45.4	-13.0	44	34.9	22.7	-12.1	7	0.8
Brickey-McCloud Elementary	63.5	61.9	-1.6	130	38.6	35.5	-3.2	19	-1.6
Carter Elementary	48.8	44.4	-4.4	99					
Cedar Bluff Elementary	60.3	56.9	-3.4	157	32.2	29.8	-2.5	26	0.9
Chilhowee Intermediate	63.3	56.6	-6.7	77	35.8	29.0	-6.8	13	-0.2
Christenberry Elementary	46.0	53.3	7.3	52	31.8	42.5	10.7	11	3.4
Copper Ridge Elementary	59.3	56.5	-2.8	90					
Corryton Elementary	50.9	42.9	-8.0	36					
Dogwood Elementary	50.3	46.4	-3.9	61	29.8	22.9	-6.9	32	-3.0
East Knox County Elementary	40.6	38.2	-2.4	77	21.0	29.0	8.0	1	
Farragut Intermediate	73.2	68.9	-4.4	310	49.2	42.9	-6.2	57	-1.9
Fountain City Elementary	59.7	50.7	-8.9	52	36.2	27.1	-9.1	10	-0.2
Gap Creek Elementary	53.0	47.2	-5.8	25					
Gibbs Elementary	66.6	58.8	-7.8	103	40.7	39.0	-1.8	24	6.1
Green Elementary	46.2	40.5	-5.7	27	25.2	18.8	-6.4	11	-0.7
Halls Elementary	59.8	51.6	-8.2	124	28.7	33.3	4.7	6	12.9
Hardin Valley Elementary	63.4	58.8	-4.7	144					
Inskip Elementary	42.7	41.9	-.8	59	26.3	30.0	3.7	3	
Karns Elementary	60.4	54.0	-6.4	160					
Lonsdale Elementary	37.9	38.0	.1	51					
Maynard Elementary	43.2	39.8	-3.4	24					
Mooreland Heights Elementary	44.9	39.7	-5.2	40	16.0	11.7	-4.3	3	
Mount Olive Elementary	58.4	52.6	-5.9	38					
New Hopewell Elementary	57.5	53.3	-4.2	26	27.4	25.5	-1.9	8	2.4
Northshore Elementary	69.8	66.8	-3.0	92	54.6	49.4	-5.2	10	-2.2
Norwood Elementary	47.3	41.9	-5.4	58	34.9	36.0	1.1	20	6.5
Pleasant Ridge Elementary	65.4	61.9	-3.5	51					
Pond Gap Elementary	66.6	58.3	-8.3	10	34.8	29.9	-5.0	26	3.3
Powell Elementary	58.0	59.0	1.0	105	36.0	33.9	-2.1	30	-3.1
Ritta Elementary	50.6	48.8	-1.8	81					
Rocky Hill Elementary	66.8	68.2	1.4	112	25.5	42.5	17.0	2	
Sarah Moore Greene	32.0	28.2	-3.8	78					
Sequoyah Elementary	79.1	74.6	-4.5	63	48.3	65.8	17.5	4	
Shannondale Elementary	63.1	59.9	-3.2	65					
South Knoxville Elementary	64.5	56.6	-7.9	10	32.4	25.3	-7.1	10	0.8
Spring Hill Elementary	72.2	65.0	-7.2	16	37.6	36.3	-1.3	25	5.9
Sterchi Elementary	61.5	62.0	.5	52					
West Haven Elementary	63.8	59.7	-4.1	25	40.3	34.4	-5.9	7	-1.8
West Hills Elementary	60.2	55.8	-4.4	81	31.2	31.6	.4	12	4.8
West View Elementary	37.6	29.1	-8.4	27					
Knox County	**59.5**	**55.6**	**-4.0**	**3558**	**37.6**	**34.4**	**-3.3**	**515**	**0.7**

Appendix B: Additional Elementary Reading Support

School		AERS Student								Differences		
		No				Yes						
		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth
		Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	Mean	Mean	Mean
Grade One	Adrian Burnett	557.7	572.1	14.4	74	535.9	509.7	-26.2	7	-21.8	-62.4	-40.6
	Amherst	582.6	572.8	-9.8	129	546.5	513.0	-33.5	1	-36.1	-59.8	-23.7
	Ball Camp	587.5	579.9	-7.6	80	557.8	535.7	-22.1	6	-29.7	-44.2	-14.5
	Bonny Kate	582.2	565.1	-17.1	48							
	Christenberry	561.9	559.5	-2.4	81							
	Copper Ridge	584.6	577.4	-7.2	74	540.5	523.0	-17.5	4	-44.2	-54.4	-10.2
	Dogwood	559.4	556.8	-2.6	108	561.7	518.0	-43.7	1	2.3	-38.8	-41.1
	Fountain City	576.9	578.1	1.2	49							
	Gibbs	567.5	566.6	-.9	119	542.8	537.7	-5.1	6	-24.7	-28.9	-4.3
	Green	560.3	550.7	-9.6	50							
	Halls	569.7	578.1	8.5	115	522.7	499.2	-23.5	5	-47.0	-78.9	-32.0
	Inskip	572.6	568.8	-3.9	84	497.1	476.0	-21.1	1	-75.5	-92.8	-17.2
	Karns	570.2	571.9	1.7	182							
	Lonsdale	555.7	543.5	-12.2	58	524.3	491.8	-32.5	6	-31.4	-51.6	-20.2
	Norwood	559.4	557.1	-2.3	86	520.3	489.7	-30.6	7	-39.1	-67.4	-28.4
	Pond Gap	564.7	547.5	-17.2	44							
	Sarah Moore Greene	558.7	556.8	-1.9	99	518.9	503.4	-15.4	9	-39.8	-53.3	-13.5
	Spring Hill	559.2	552.2	-7.0	77							
	West Haven	574.0	569.9	-4.1	54	537.9	533.6	-4.3	5	-36.0	-36.3	-0.2
	Total	568.8	565.8	-3.0	1611	532.5	511.7	-20.7	58	-36.3	-54.1	-17.7
Grade Two	Adrian Burnett	596.8	589.3	-7.6	91	581.1	557.5	-23.6	6	-15.7	-31.8	-16.0
	Amherst	613.6	615.1	1.5	120							
	Ball Camp	616.5	622.2	5.7	92	588.5	606.9	18.3	7	-28.0	-15.3	12.7
	Bonny Kate	605.0	606.1	1.1	67							
	Christenberry	589.8	593.5	3.7	72							
	Copper Ridge	601.3	599.9	-1.4	79							
	Dogwood	594.2	588.6	-5.6	96	579.8	587.5	7.7	2	-14.4	-1.1	13.3
	Fountain City	605.8	603.8	-1.9	50							
	Gibbs	602.0	614.3	12.4	121							
	Green	584.0	580.4	-3.6	44	565.1	553.7	-11.4	7	-18.9	-26.7	-7.8
	Halls	612.2	608.1	-4.1	101	566.4	551.4	-15.0	5	-45.8	-56.7	-10.9
	Inskip	610.5	608.1	-2.5	72							
	Karns	606.3	609.4	3.1	165							
	Lonsdale	583.6	574.5	-9.1	71							
	Norwood	597.1	600.7	3.6	84							
	Pond Gap	593.1	596.3	3.2	38							
	Sarah Moore Greene	588.9	585.0	-3.8	82	554.2	544.5	-9.6	13	-34.7	-40.5	-5.8
	Spring Hill	600.4	594.0	-6.3	72							
	West Haven	613.5	604.9	-8.6	49	570.1	564.0	-6.1	3	-43.4	-40.9	2.5
	Total	601.9	601.5	-.3	1566	569.0	562.1	-6.9	43	-32.9	-39.4	-6.5

School		AERS Student								Differences		
		No				Yes						
		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth
		Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	Mean	Mean	Mean
Grade Three	Adrian Burnett	737.1	736.2	-.9	96	715.9	706.7	-9.2	11	-21.2	-29.5	-8.3
	Amherst	754.7	753.3	-1.4	101	725.7	708.5	-17.2	11	-29.0	-44.8	-15.8
	Ball Camp	756.4	760.2	3.7	82	732.6	733.8	1.3	6	-23.9	-26.3	-2.5
	Bonny Kate	757.3	762.1	4.8	44	734.4	725.3	-9.2	12	-22.9	-36.8	-13.9
	Chilhowee Intermediate	744.4	749.7	5.3	65	735.8	728.1	-7.7	14	-8.6	-21.7	-13.0
	Christenberry	734.1	744.0	9.9	65	733.5	736.9	3.5	15	-0.6	-7.1	-6.4
	Copper Ridge	748.6	739.5	-9.1	89	751.3	735.3	-15.9	6	2.6	-4.2	-6.8
	Dogwood	735.6	731.3	-4.3	113							
	Fountain City	747.5	742.3	-5.2	41	732.1	721.1	-11.0	9	-15.4	-21.2	-5.8
	Gibbs	751.3	750.5	-.9	111	741.5	744.8	3.3	6	-9.9	-5.6	4.2
	Green	733.8	716.3	-17.5	35	705.7	692.9	-12.8	7	-28.1	-23.4	4.7
	Halls	759.8	757.9	-1.9	89	720.6	713.3	-7.3	15	-39.2	-44.6	-5.4
	Inskip	752.6	753.9	1.3	36	719.4	720.8	1.4	33	-33.3	-33.1	0.1
	Karns	758.2	760.9	2.8	162	742.8	730.0	-12.8	19	-15.4	-30.9	-15.6
	Lonsdale	725.6	724.5	-1.2	40	758.9	761.8	3.0	6	33.2	37.4	4.1
	Norwood	737.3	735.4	-2.0	88	742.8	734.3	-8.5	8	5.5	-1.1	-6.6
	Pond Gap	741.8	743.7	1.9	45	723.9	722.4	-1.5	10	-18.0	-21.3	-3.3
	Sarah Moore Greene	723.5	715.6	-7.9	75	702.4	687.7	-14.7	11	-21.1	-27.9	-6.8
	Spring Hill	749.7	747.1	-2.6	32	735.1	715.3	-19.8	14	-14.6	-31.8	-17.2
	West Haven	738.4	740.2	1.7	46	731.4	732.7	1.3	7	-7.1	-7.5	-0.4
	Total	745.7	744.8	-.9	1455	729.0	722.2	-6.9	220	-16.7	-22.6	-5.9
Grade Four	Adrian Burnett	49.1	48.8	-.3	71	38.2	32.8	-5.4	11	-10.9	-16.0	-5.0
	Amherst	62.5	56.1	-6.4	98	39.4	38.5	-.8	11	-23.1	-17.5	5.6
	Ball Camp	62.7	63.9	1.2	70	46.3	52.5	6.2	6	-16.4	-11.4	5.0
	Bonny Kate	57.2	56.9	-.3	40	50.9	53.3	2.4	7	-6.4	-3.6	2.8
	Chilhowee Intermediate	53.9	53.7	-.2	69	41.3	42.8	1.5	6	-12.5	-10.9	1.7
	Christenberry	56.1	56.3	.1	54	46.9	32.4	-14.4	7	-9.3	-23.8	-14.6
	Copper Ridge	58.2	62.0	3.9	71	35.8	28.5	-7.3	6	-22.3	-33.5	-11.2
	Dogwood	44.1	39.4	-4.8	93	34.4	26.9	-7.5	8	-9.8	-12.5	-2.7
	Fountain City	53.7	48.0	-5.7	73							
	Gibbs	57.0	56.9	-.1	116	46.3	50.4	4.1	7	-10.7	-6.5	4.3
	Green	37.2	35.1	-2.1	39	12.4	13.8	1.4	5	-24.8	-21.3	3.5
	Halls	55.5	53.2	-2.3	138	27.2	25.1	-2.1	15	-28.3	-28.1	0.2
	Inskip	47.4	44.6	-2.8	68							
	Karns	58.4	61.5	3.1	196	43.3	34.0	-9.3	13	-15.1	-27.5	-12.4
	Lonsdale	39.7	34.9	-4.8	32	32.3	22.8	-9.5	6	-7.4	-12.0	-4.7
	Norwood	44.7	45.5	.8	74	26.1	27.8	1.6	8	-18.6	-17.7	0.9
	Pond Gap	42.3	40.0	-2.3	40	24.8	21.7	-3.2	6	-17.5	-18.4	-0.9
	Sarah Moore Greene	38.1	33.0	-5.2	96	22.9	16.2	-6.7	15	-15.2	-16.8	-1.6
	Spring Hill	49.0	47.3	-1.8	47	40.7	37.9	-2.7	15	-8.4	-9.3	-1.0
	West Haven	58.7	57.6	-1.0	46	30.6	27.8	-2.8	5	-28.1	-29.8	-1.8
	Total	52.6	51.3	-1.3	1531	35.5	32.0	-3.5	157	-17.1	-19.3	-2.2

School		AERS Student								Differences		
		No				Yes						
		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth		Predicted Score	Observed Score	Growth
		Mean	Mean	Mean	Count	Mean	Mean	Mean	Count	Mean	Mean	Mean
Grade Five	Adrian Burnett	45.9	41.5	-4.4	97	31.2	30.0	-1.2	6	-14.8	-11.5	3.2
	Amherst	58.1	51.0	-7.0	118	31.6	30.9	-.7	10	-26.5	-20.1	6.3
	Ball Camp	58.2	54.0	-4.2	72	42.4	36.6	-5.8	5	-15.8	-17.4	-1.6
	Bonny Kate	59.2	46.4	-12.8	43	33.1	20.0	-13.1	8	-26.1	-26.4	-0.3
	Chilhowee Intermediate	61.3	54.8	-6.5	83	35.1	26.4	-8.7	7	-26.2	-28.4	-2.2
	Christenberry	45.3	52.4	7.1	57	26.3	42.3	16.0	6	-19.0	-10.1	8.9
	Copper Ridge	60.5	58.1	-2.4	84	42.7	34.2	-8.5	6	-17.8	-23.9	-6.1
	Dogwood	43.5	38.8	-4.8	86	40.1	33.1	-7.0	7	-3.4	-5.6	-2.2
	Fountain City	55.9	46.9	-9.0	62							
	Gibbs	62.3	55.3	-7.0	121	49.3	49.5	.2	6	-13.0	-5.8	7.2
	Green	42.8	36.9	-5.9	34	17.0	11.3	-5.8	4	-25.8	-25.7	0.2
	Halls	61.1	52.6	-8.5	118	31.3	32.5	1.3	12	-29.9	-20.1	9.8
	Inskip	42.2	41.8	-.5	61	20.0	12.0	-8.0	1	-22.2	-29.8	-7.5
	Karns	62.1	56.0	-6.1	147	41.4	31.7	-9.7	13	-20.7	-24.3	-3.6
	Lonsdale	37.9	38.0	.1	51							
	Norwood	45.1	40.9	-4.2	69	36.4	36.6	.1	9	-8.7	-4.3	4.4
	Pond Gap	43.8	37.9	-5.9	34	41.0	35.0	-6.0	2	-2.8	-2.9	-0.1
	Sarah Moore Greene	33.3	29.5	-3.8	69	21.8	18.3	-3.4	9	-11.6	-11.2	0.4
	Spring Hill	57.3	52.4	-4.9	30	34.4	34.3	-.1	11	-22.9	-18.1	4.8
	West Haven	62.3	57.1	-5.2	28	33.5	34.0	.5	4	-28.8	-23.1	5.7
	Total	**53.4**	**48.3**	**-5.1**	**1464**	**34.4**	**31.3**	**-3.1**	**126**	**-19.0**	**-17.1**	**1.9**

Appendix C.1: Staffing Ratios

		Number of Classes															
		1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students		Total Number of Classes	
		Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %
Austin-East High	Algebra I	1	8.3%	2	16.7%	7	58.3%	2	16.7%							12	100.0%
	Algebra II			3	33.3%	5	55.6%	1	11.1%							9	100.0%
	Biology I	2	18.2%	2	18.2%	3	27.3%	4	36.4%							11	100.0%
	Chemistry I			3	30.0%	4	40.0%	2	20.0%	1	10.0%					10	100.0%
	English I	1	9.1%	1	9.1%	6	54.5%	3	27.3%							11	100.0%
	English II			1	12.5%	2	25.0%	3	37.5%	2	25.0%					8	100.0%
	English III			5	50.0%	3	30.0%	2	20.0%							10	100.0%
	US History	1	11.1%	4	44.4%	3	33.3%	1	11.1%							9	100.0%
	Total	5	6.3%	21	26.3%	33	41.3%	18	22.5%	3	3.8%					80	100.0%
Bearden High	Algebra I					4	21.1%	7	36.8%	3	15.8%	4	21.1%	1	5.3%	19	100.0%
	Algebra II			1	5.3%	2	10.5%	5	26.3%	7	36.8%	4	21.1%			19	100.0%
	Biology I	1	4.5%					2	9.1%	9	40.9%	10	45.5%			22	100.0%
	Chemistry I					5	20.0%	7	28.0%	4	16.0%	4	16.0%	5	20.0%	25	100.0%
	English I	1	4.8%			1	4.8%	2	9.5%	6	28.6%	7	33.3%	4	19.0%	21	100.0%
	English II			1	5.0%	1	5.0%	3	15.0%	8	40.0%	3	15.0%	4	20.0%	20	100.0%
	English III	1	7.1%	1	7.1%			1	7.1%	5	35.7%	3	21.4%	3	21.4%	14	100.0%
	US History							1	12.5%			5	62.5%	2	25.0%	8	100.0%
	Total	3	2.0%	3	2.0%	13	8.8%	28	18.9%	42	28.4%	40	27.0%	19	12.8%	148	100.0%
Carter High	Algebra I					6	42.9%	3	21.4%	4	28.6%	1	7.1%			14	100.0%
	Algebra II			1	10.0%	1	10.0%	3	30.0%	4	40.0%	1	10.0%			10	100.0%
	Biology I					1	9.1%	4	36.4%	5	45.5%	1	9.1%			11	100.0%
	Chemistry I							7	70.0%	1	10.0%	1	10.0%	1	10.0%	10	100.0%
	English I			1	7.7%	2	15.4%	5	38.5%	4	30.8%	1	7.7%			13	100.0%
	English II			1	8.3%	1	8.3%	3	25.0%	5	41.7%	2	16.7%			12	100.0%
	English III			4	33.3%	2	16.7%	3	25.0%	3	25.0%					12	100.0%
	US History					1	12.5%	2	25.0%	2	25.0%	3	37.5%			8	100.0%
	Total			7	7.8%	14	15.6%	30	33.3%	28	31.1%	10	11.1%	1	1.1%	90	100.0%
Central High	Algebra I			2	13.3%	2	13.3%	3	20.0%	8	53.3%					15	100.0%
	Algebra II					1	8.3%	4	33.3%	6	50.0%	1	8.3%			12	100.0%
	Biology I							4	28.6%	9	64.3%	1	7.1%			14	100.0%
	Chemistry I					4	30.8%	6	46.2%	3	23.1%					13	100.0%
	English I			1	6.7%			7	46.7%	5	33.3%	2	13.3%			15	100.0%
	English II							3	25.0%	4	33.3%	4	33.3%	1	8.3%	12	100.0%
	English III	1	9.1%	1	9.1%			4	36.4%	3	27.3%	1	9.1%	1	9.1%	11	100.0%
	US History					1	12.5%			2	25.0%	4	50.0%	1	12.5%	8	100.0%
	Total	1	1.0%	4	4.0%	8	8.0%	31	31.0%	40	40.0%	13	13.0%	3	3.0%	100	100.0%
Farragut High	Algebra I			1	7.1%			4	28.6%	8	57.1%	1	7.1%			14	100.0%
	Algebra II							5	27.8%	8	44.4%	5	27.8%			18	100.0%
	Biology I							7	38.9%	6	33.3%	5	27.8%			18	100.0%
	Chemistry I							2	11.8%	11	64.7%	4	23.5%			17	100.0%
	English I							2	12.5%	6	37.5%	8	50.0%			16	100.0%
	English II			1	5.3%			6	31.6%	6	31.6%	5	26.3%	1	5.3%	19	100.0%
	English III	1	7.7%	1	7.7%	1	7.7%	4	30.8%	3	23.1%	3	23.1%			13	100.0%
	US History					1	9.1%	4	36.4%	1	9.1%	4	36.4%	1	9.1%	11	100.0%
	Total	1	.8%	3	2.4%	2	1.6%	34	27.0%	49	38.9%	35	27.8%	2	1.6%	126	100.0%

		Number of Classes															
		1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students		Total Number of Classes	
		Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %
Fulton High	Algebra I			3	20.0%	1	6.7%	7	46.7%	3	20.0%	1	6.7%			15	100.0%
	Algebra II			1	9.1%	4	36.4%	3	27.3%	3	27.3%					11	100.0%
	Biology I	1	16.7%	1	16.7%	3	50.0%	1	16.7%							6	100.0%
	Chemistry I					8	66.7%	3	25.0%	1	8.3%					12	100.0%
	English I	2	12.5%	2	12.5%	2	12.5%	5	31.3%	5	31.3%					16	100.0%
	English II	2	12.5%	2	12.5%	4	25.0%	6	37.5%	2	12.5%					16	100.0%
	English III	4	28.6%	2	14.3%	4	28.6%	2	14.3%	2	14.3%					14	100.0%
	US History					1	11.1%	4	44.4%	4	44.4%					9	100.0%
	Total	9	9.1%	11	11.1%	27	27.3%	31	31.3%	20	20.2%	1	1.0%			99	100.0%
Gibbs High	Algebra I			4	36.4%			3	27.3%	2	18.2%	2	18.2%			11	100.0%
	Algebra II					1	9.1%			4	36.4%	5	45.5%	1	9.1%	11	100.0%
	Biology I			1	6.7%					5	33.3%	7	46.7%	2	13.3%	15	100.0%
	Chemistry I					1	12.5%	5	62.5%	2	25.0%					8	100.0%
	English I			1	8.3%	2	16.7%	4	33.3%			3	25.0%	2	16.7%	12	100.0%
	English II							1	9.1%	6	54.5%	2	18.2%	2	18.2%	11	100.0%
	English III	1	10.0%					1	10.0%	3	30.0%	4	40.0%	1	10.0%	10	100.0%
	US History					1	11.1%	1	11.1%	7	77.8%					9	100.0%
	Total	1	1.1%	6	6.9%	5	5.7%	15	17.2%	29	33.3%	23	26.4%	8	9.2%	87	100.0%
Halls High	Algebra I			4	25.0%	3	18.8%	5	31.3%	3	18.8%	1	6.3%			16	100.0%
	Algebra II					3	27.3%	2	18.2%	2	18.2%	4	36.4%			11	100.0%
	Biology I					1	6.7%	5	33.3%	6	40.0%	3	20.0%			15	100.0%
	Chemistry I	1	6.7%			3	20.0%	6	40.0%	3	20.0%	2	13.3%			15	100.0%
	English I			3	17.6%	1	5.9%	6	35.3%	5	29.4%	1	5.9%	1	5.9%	17	100.0%
	English II			1	7.1%	1	7.1%	3	21.4%	7	50.0%	2	14.3%			14	100.0%
	English III	1	7.7%					5	38.5%	4	30.8%	3	23.1%			13	100.0%
	US History					1	9.1%	4	36.4%	4	36.4%	1	9.1%	1	9.1%	11	100.0%
	Total	2	1.8%	8	7.1%	13	11.6%	36	32.1%	34	30.4%	17	15.2%	2	1.8%	112	100.0%
Hardin Valley Academy	Algebra I			1	5.0%	5	25.0%	2	10.0%	2	10.0%	4	20.0%	6	30.0%	20	100.0%
	Algebra II			2	11.1%					5	27.8%	7	38.9%	4	22.2%	18	100.0%
	Biology I			1	6.3%	1	6.3%	1	6.3%	2	12.5%	4	25.0%	7	43.8%	16	100.0%
	Chemistry I							1	6.3%	1	6.3%	8	50.0%	6	37.5%	16	100.0%
	English I			6	20.7%	9	31.0%	4	13.8%	3	10.3%	4	13.8%	3	10.3%	29	100.0%
	English II					3	15.8%	3	15.8%	4	21.1%	1	5.3%	8	42.1%	19	100.0%
	English III			1	7.1%	3	21.4%	1	7.1%	3	21.4%	2	14.3%	4	28.6%	14	100.0%
	US History									2	20.0%	2	20.0%	6	60.0%	10	100.0%
	Total			11	7.7%	21	14.8%	12	8.5%	22	15.5%	32	22.5%	44	31.0%	142	100.0%
Karns High	Algebra I			2	12.5%	3	18.8%	8	50.0%	3	18.8%					16	100.0%
	Algebra II			1	6.7%	3	20.0%	6	40.0%	4	26.7%	1	6.7%			15	100.0%
	Biology I	1	6.7%			1	6.7%	3	20.0%	7	46.7%	3	20.0%			15	100.0%
	Chemistry I			1	7.7%	3	23.1%	6	46.2%	3	23.1%					13	100.0%
	English I					2	14.3%	5	35.7%	4	28.6%	3	21.4%			14	100.0%
	English II			1	6.3%	2	12.5%	4	25.0%	8	50.0%	1	6.3%			16	100.0%
	English III			2	20.0%	1	10.0%	1	10.0%	3	30.0%	3	30.0%			10	100.0%
	US History									5	55.6%	4	44.4%			9	100.0%
	Total	1	.9%	7	6.5%	15	13.9%	33	30.6%	37	34.3%	15	13.9%			108	100.0%

		Number of Classes															
		1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students		21 to 25 Students		26 to 30 Students		More than 30 Students		Total Number of Classes	
		Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %	Count	Row N %
L & N STEM Academy	Algebra I	3	42.9%	1	14.3%			1	14.3%	1	14.3%	1	14.3%			7	100.0%
	Algebra II							1	20.0%	2	40.0%	2	40.0%			5	100.0%
	Biology I					1	25.0%			1	25.0%	1	25.0%	1	25.0%	4	100.0%
	Chemistry I							1	12.5%	1	12.5%	3	37.5%	3	37.5%	8	100.0%
	English I			1	16.7%					1	16.7%	1	16.7%	3	50.0%	6	100.0%
	English II					2	28.6%	1	14.3%	1	14.3%	3	42.9%			7	100.0%
	English III									2	66.7%	1	33.3%			3	100.0%
	US History									1	33.3%			2	66.7%	3	100.0%
	Total	3	7.0%	2	4.7%	3	7.0%	4	9.3%	10	23.3%	12	27.9%	9	20.9%	43	100.0%
Powell High	Algebra I					4	22.2%	7	38.9%	5	27.8%	1	5.6%	1	5.6%	18	100.0%
	Algebra II	1	6.3%			5	31.3%	3	18.8%	2	12.5%	3	18.8%	2	12.5%	16	100.0%
	Biology I			2	10.0%	4	20.0%	4	20.0%	5	25.0%	3	15.0%	2	10.0%	20	100.0%
	Chemistry I			2	11.8%			7	41.2%	5	29.4%	3	17.6%			17	100.0%
	English I									3	23.1%	6	46.2%	4	30.8%	13	100.0%
	English II							1	7.7%	3	23.1%	6	46.2%	3	23.1%	13	100.0%
	English III					1	8.3%	1	8.3%	5	41.7%	4	33.3%	1	8.3%	12	100.0%
	US History							2	20.0%	3	30.0%	3	30.0%	2	20.0%	10	100.0%
	Total	1	.8%	4	3.4%	14	11.8%	25	21.0%	31	26.1%	29	24.4%	15	12.6%	119	100.0%
South-Doyle High	Algebra I	2	10.5%			5	26.3%	5	26.3%	6	31.6%	1	5.3%			19	100.0%
	Algebra II					2	16.7%	1	8.3%	5	41.7%	3	25.0%	1	8.3%	12	100.0%
	Biology I			2	12.5%			4	25.0%	5	31.3%	4	25.0%	1	6.3%	16	100.0%
	Chemistry I					1	6.7%	3	20.0%	7	46.7%	2	13.3%	2	13.3%	15	100.0%
	English I			3	20.0%			2	13.3%	6	40.0%	3	20.0%	1	6.7%	15	100.0%
	English II			3	21.4%	1	7.1%	2	14.3%	2	14.3%	5	35.7%	1	7.1%	14	100.0%
	English III			2	16.7%	2	16.7%	4	33.3%	2	16.7%	2	16.7%			12	100.0%
	US History									2	28.6%	5	71.4%			7	100.0%
	Total	2	1.8%	10	9.1%	11	10.0%	21	19.1%	35	31.8%	25	22.7%	6	5.5%	110	100.0%
West High	Algebra I	3	13.0%	1	4.3%	2	8.7%	11	47.8%	6	26.1%					23	100.0%
	Algebra II					3	23.1%	4	30.8%	2	15.4%	4	30.8%			13	100.0%
	Biology I	1	6.7%					7	46.7%	2	13.3%	5	33.3%			15	100.0%
	Chemistry I					2	14.3%	5	35.7%	5	35.7%	2	14.3%			14	100.0%
	English I	2	11.1%	2	11.1%	3	16.7%	4	22.2%	4	22.2%	1	5.6%	2	11.1%	18	100.0%
	English II			1	7.7%	3	23.1%	4	30.8%	3	23.1%	1	7.7%	1	7.7%	13	100.0%
	English III	1	11.1%			3	33.3%			3	33.3%	2	22.2%			9	100.0%
	US History					3	30.0%			3	30.0%	3	30.0%	1	10.0%	10	100.0%
	Total	7	6.1%	4	3.5%	19	16.5%	35	30.4%	28	24.3%	18	15.7%	4	3.5%	115	100.0%
Total	Algebra I	9	4.1%	21	9.6%	42	19.2%	68	31.1%	54	24.7%	17	7.8%	8	3.7%	219	100.0%
	Algebra II	1	.6%	9	5.0%	30	16.7%	38	21.1%	54	30.0%	40	22.2%	8	4.4%	180	100.0%
	Biology I	6	3.0%	9	4.5%	15	7.6%	46	23.2%	62	31.3%	47	23.7%	13	6.6%	198	100.0%
	Chemistry I	1	.5%	6	3.1%	31	16.1%	61	31.6%	48	24.9%	29	15.0%	17	8.8%	193	100.0%
	English I	6	2.8%	21	9.7%	28	13.0%	49	22.7%	52	24.1%	40	18.5%	20	9.3%	216	100.0%
	English II	2	1.0%	12	6.2%	20	10.3%	43	22.2%	61	31.4%	35	18.0%	21	10.8%	194	100.0%
	English III	10	6.4%	19	12.1%	20	12.7%	29	18.5%	41	26.1%	28	17.8%	10	6.4%	157	100.0%
	US History	1	.8%	4	3.3%	12	9.8%	19	15.6%	36	29.5%	34	27.9%	16	13.1%	122	100.0%
	Total	36	2.4%	101	6.8%	198	13.4%	353	23.9%	408	27.6%	270	18.3%	113	7.6%	1479	100.0%

Appendix C.2: Staffing Ratios

	Subject							
	Algebra I		Algebra II		Biology I		Chemistry I	
	Growth		Growth		Growth		Growth	
	Mean	Count	Mean	Count	Mean	Count	Mean	Count
Austin-East High	33.2	124	29.5	98	6.3	109	-22.8	110
Bearden High	-6.7	350	13.8	359	-.3	475	9.3	503
Carter High	21.0	244	7.7	191	4.9	213	9.0	197
Central High	-14.0	239	-1.0	224	4.5	283	20.2	206
Farragut High	4.3	259	16.9	359	3.8	364	14.2	341
Fulton High	-9.7	221	.6	166	-15.6	61	-.4	156
Gibbs High	-.3	175	4.7	239	-2.9	349	8.0	132
Halls High	16.9	247	20.3	221	5.5	317	17.6	266
Hardin Valley Academy	-6.1	404	13.7	386	2.3	371	11.6	403
Karns High	-11.6	250	2.0	244	2.2	303	-1.5	192
L & N STEM Academy	-18.3	70	-7.2	105	-13.0	84	.2	194
Powell High	8.2	327	11.3	286	-1.0	379	15.1	318
South-Doyle High	6.4	307	15.5	257	6.6	306	4.3	336
West High	-6.3	318	-9.4	216	-8.9	259	22.4	223
Total	.7	3535	9.3	3351	1.0	3873	9.6	3577

	Subject									
	English I		English II		English III		US History		Total	
	Growth		Growth		Growth		Growth		Growth	
	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count
Austin-East High	5.4	127	10.5	107	11.4	107	2.8	85	9.8	867
Bearden High	-1.9	472	-1.9	440	6.0	275	9.5	208	2.9	3082
Carter High	2.3	231	.0	227	5.3	171	9.0	169	7.5	1643
Central High	-1.9	282	1.0	280	4.7	198	-2.5	180	1.1	1892
Farragut High	.3	365	1.5	353	7.4	212	6.8	220	6.9	2473
Fulton High	-2.8	221	2.0	203	8.9	151	6.4	155	-.7	1334
Gibbs High	-4.4	241	.4	256	7.2	206	7.2	181	1.6	1779
Halls High	-2.5	311	2.3	280	4.2	252	7.7	234	8.4	2128
Hardin Valley Academy	-1.0	462	-.2	409	4.2	284	4.1	265	3.4	2984
Karns High	-3.1	275	1.8	297	1.8	178	2.5	205	-.8	1944
L & N STEM Academy	-1.0	129	-6.0	128	5.9	59	5.4	71	-3.9	840
Powell High	-2.6	356	-1.4	326	3.1	259	6.6	232	4.6	2483
South-Doyle High	1.5	285	2.9	281	2.0	199	7.8	171	5.8	2142
West High	-.9	276	5.2	209	4.5	139	7.5	192	.8	1832
Total	-1.2	4033	.9	3796	5.2	2690	5.9	2568	3.6	27423

Appendix C.3: Staffing Ratios

		Class Size							
		1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students	
		Growth		Growth		Growth		Growth	
		Mean	Count	Mean	Count	Mean	Count	Mean	Count
Subject	Algebra I	-26.1	15	-.9	160	.5	472	1.8	1106
	Algebra II	-89.7	2	.5	64	7.2	359	6.5	637
	Biology I	-6.1	19	-14.8	64	-7.1	181	3.8	755
	Chemistry I	50.6	5	8.2	43	7.1	369	8.1	979
	English I	-18.2	12	-16.1	152	-1.1	322	-.5	809
	English II	-35.9	3	-5.8	81	1.4	242	1.6	711
	English III	-20.9	26	-6.6	145	9.4	237	4.4	466
	US History	-3.4	2	6.5	33	5.0	145	5.8	307
	Total	-15.6	84	-5.9	742	3.1	2327	3.7	5770

		Class Size							
		21 to 25 Students		26 to 30 Students		More than 30 Students		Total	
		Growth		Growth		Growth		Growth	
		Mean	Count	Mean	Count	Mean	Count	Mean	Count
Subject	Algebra I	.7	1124	.5	422	-.4	236	.7	3535
	Algebra II	11.5	1096	9.2	963	13.3	230	9.3	3351
	Biology I	2.1	1306	-.5	1172	2.8	376	1.0	3873
	Chemistry I	9.6	978	13.3	694	8.8	509	9.6	3577
	English I	.7	1106	-1.2	1028	-1.2	604	-1.2	4033
	English II	.3	1256	1.0	880	1.8	623	.9	3796
	English III	6.5	828	5.4	696	6.5	292	5.2	2690
	US History	6.3	761	6.4	864	4.7	456	5.9	2568
	Total	4.4	8455	4.0	6719	3.9	3326	3.6	27423

Appendix C.4: Staffing Ratios

Class Size Growth Minus Subject Growth		Subject Growth Means	Class Size							
			1 to 5 Students		6 to 10 Students		11 to 15 Students		16 to 20 Students	
			Growth		Growth		Growth		Growth	
			Mean	Count	Mean	Count	Mean	Count	Mean	Count
Subject	Algebra I	.7	-26.8	15	-1.7	160	-0.2	472	1.1	1106
	Algebra II	9.3	-99.0	2	-8.7	64	-2.1	359	-2.7	637
	Biology I	1.0	-7.1	19	-15.7	64	-8.1	181	2.8	755
	Chemistry I	9.6	41.0	5	-1.4	43	-2.5	369	-1.5	979
	English I	-1.2	-17.1	12	-14.9	152	0.1	322	0.6	809
	English II	.9	-36.8	3	-6.7	81	0.5	242	0.8	711
	English III	5.2	-26.1	26	-11.7	145	4.3	237	-0.8	466
	US History	5.9	-9.3	2	0.6	33	-0.9	145	-0.1	307
	Total	3.6	-19.2	84	-9.5	742	-0.6	2327	0.1	5770

Class Size Growth Minus Subject Growth		Subject Growth Means	Class Size							
			21 to 25 Students		26 to 30 Students		More than 30 Students		Total	
			Growth		Growth		Growth		Growth	
			Mean	Count	Mean	Count	Mean	Count	Mean	Count
Subject	Algebra I	.7	0.0	1124	-0.2	422	-1.1	236	0.0	3535
	Algebra II	9.3	2.2	1096	-0.1	963	4.0	230	0.0	3351
	Biology I	1.0	1.2	1306	-1.4	1172	1.8	376	0.0	3873
	Chemistry I	9.6	0.0	978	3.7	694	-0.8	509	0.0	3577
	English I	-1.2	1.8	1106	-0.1	1028	0.0	604	0.0	4033
	English II	.9	-0.5	1256	0.1	880	1.0	623	0.0	3796
	English III	5.2	1.4	828	0.3	696	1.3	292	0.0	2690
	US History	5.9	0.4	761	0.5	864	-1.2	456	0.0	2568
	Total	3.6	0.8	8455	0.3	6719	0.3	3326	0.0	27423

Appendix D.1: Report to Mayor Burchett on the Early Literacy Program

Knox County Schools

REPORT TO MAYOR BURCHETT ON THE EARLY LITERACY PROGRAM

January, 2015

REPORT TO MAYOR BURCHETT ON THE EARLY LITERACY PROGRAM

Executive Summary

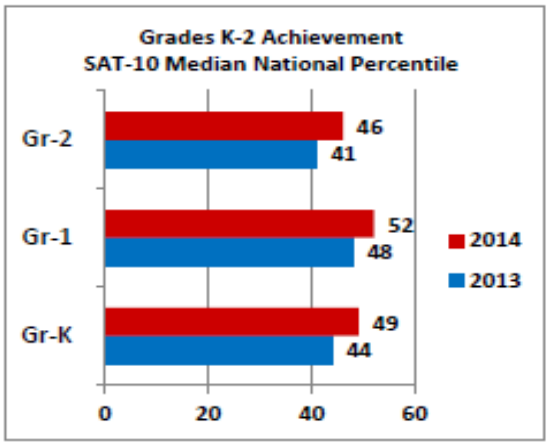
Starting with the fiscal year 2013 (FY13) budget, the Knox County Commission, at the urging of Mayor Tim Burchett, appropriated an additional $2.87 million to specifically support elementary level (grades K-5) literacy initiatives in the Knox County Schools. This investment made it possible for the Knox County Schools (KCS) to provide additional instructional support aimed at enhancing student learning and success in reading and English language arts.

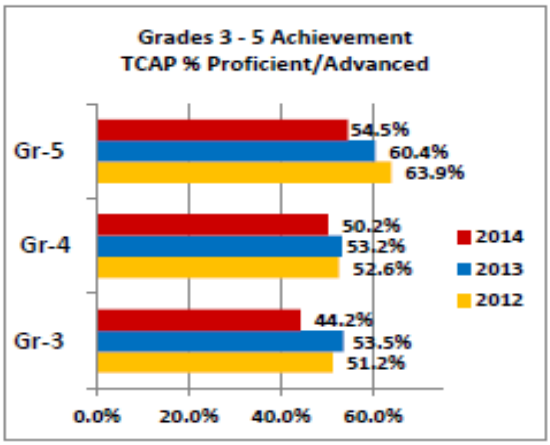
This Report to Mayor Burchett on the Early Literacy Program summarizes the FY13 and FY14 expenditures and related results. This report is a follow-up to the 2014 Educational Return on Investment (EROI) Report, published and distributed to all School Board Members, County Commissioners, and Mayor Burchett in April 2014, which included a detailed program evaluation of all funding and activities related to the early literacy program. Significant adjustments to literacy programming have been made based on this detailed program evaluation (see *Continuous Improvement* section below).

Academic Results

The academic results for the early literacy initiatives in grades K – 5 are measured based on student academic growth and achievement. **Leading indications are that the literacy program is having a positive effect in the primary grades where much of our effort has been focused, whereas the results in later grades are somewhat mixed.** In grades K, 1 and 2, the district saw significant gains in reading and language achievement, as well as steady student growth in language. In grades 3, 4, and 5 the trends have been inconsistent in both achievement and growth for the district. **The promising leading indicators in the primary grades, coupled with modifications to literacy programming made based on rigorous program evaluation, suggest that continued investment in early literacy would be both reasonable and beneficial.** The summary academic results are as follows:

Student Achievement Results





Student Growth Results

	2013 One-Year TVAAS Growth					2014 One-Year TVAAS Growth				
	Grade									
(Subject)	1	2	3	4	5	1	2	3	4	5
Reading	4.0	2.3				-2.7	-1.5			
Language	11.1	4.2				1.3	3.2			
Reading/Language Arts			3.3	2.1	0.0			0.3	-1.2	-4.1

Investment Analysis

The $2.87 million of additional funding was primarily allocated to provide teacher support personnel, including instructional coaches and instructional assistants, intervention and instructional materials, and professional development stipends and substitutes. The summary of these expenditures are as follows:

Early Literacy Initiative Expenditures	FY13	FY14
Instructional Coaches	$ 2,195,408	$ 1,801,202
Instructional Assistants	$ 316,785	$ 397,348
Other Personnel	$ 110,401	$ 115,914
Professional Development (Subs & Stipends)	$ -	$ 68,419
Intervention Materials	$ 200,000	$ 277,910
Other Materials	$ 44,904	$ 118,836
Carry-over (Available for FY15)	$ -	$ 90,371
Total	$ 2,867,498	$ 2,870,000

Management Report

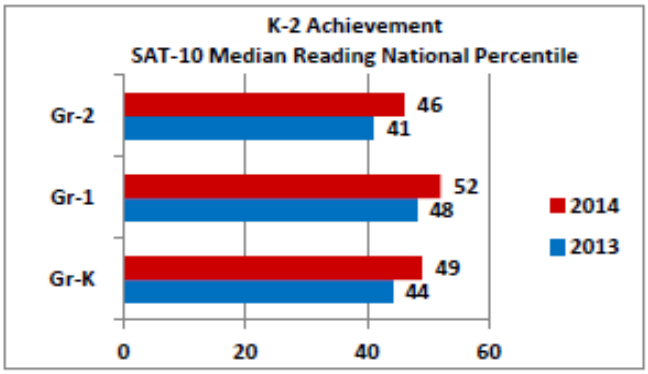
Academic Results

The academic results for the early literacy initiatives in grades K – 5 are measured based on growth and achievement. The academic results for individual grade bands are discussed below. (See Appendix A for school level *2010-2014 Elementary Reading Trends*.)

Grades K through 2

The results for kindergarten, first grade and second grade are based on the Stanford Achievement Test, Series 10(SAT-10). This early grades assessment was administered district-wide for the first time in 2012-13. The SAT-10 is a norm-referenced assessment used across the nation. The SAT-10 was used in both years of this program evaluation, and there were no changes to the test form. Therefore, this measurement was stable over the duration of the evaluation. In November, the School Board voted to discontinue administration of the SAT-10 assessment effective immediately. Therefore, SAT-10 results will not be available for the 2014-15 school year and beyond. It is unclear if a comparable metric can be used for summative assessment and program evaluation for the early grades moving forward.

- **Achievement** – For all three grades, the district overall improved in terms of mean scale score and median national percentile rank on SAT-10 from 2013 to 2014.

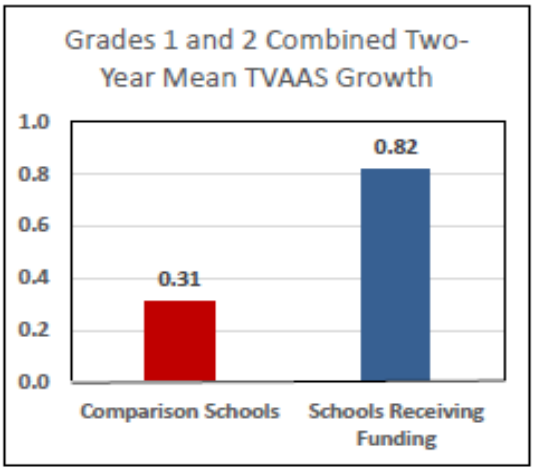


- **Growth** – In 2013, the district made positive gains in reading and language grades 1 and 2, with students outperforming predicted results based on the TVAAS model. In 2014, grades 1 and 2 did not show positive growth in reading, though language remained positive as depicted in the adjacent table.

	2013 One-Year TVAAS Growth		2014 One-Year TVAAS Growth	
Subject	Grade 1	Grade 2	Grade 1	Grade 2
Reading	4.0	2.3	-2.7	-1.5
Language	11.1	4.2	1.3	3.2

- **Growth Versus Comparison Schools** – The schools directly receiving early literacy supplemental funding in both years (2012-13 and 2013-14) performed better than those that did not in terms of growth, as detailed in the tables below .

Early Literacy Schools Funded Both Years
Adrian Burnett; Amherst; Ball Camp; Beaumont; Belle Morris; Blue Grass; Bonny Kate; Cedar Bluff; Chilhowee; Christenberry; Copper Ridge; Dogwood; East Knox; Farragut Primary; Fountain City; Gibbs; Green; Halls; Inskip; Karns; Lonsdale; Maynard; Mount Olive; New Hopewell; Norwood; Pond Gap; Powell; Sarah Moore Greene; Spring Hill; Sunnyview Primary; West Haven; West Hills
Comparison Schools Not Funded Either Year
A. L. Lotts; Brickey-McCloud; Farragut Intermediate; Gap Creek; Hardin Valley; K.A.E.C.; Mooreland Heights; Northshore; Pleasant Ridge; Ridgedale; Ritta; Shannondale; South Knoxville; West View



Grade 3

The third grade results are based on the Tennessee Comprehensive Assessment Program reading / English language arts assessment. The TCAP is a state-specific assessment that begins in 3rd grade and is used only in Tennessee.

The Tennessee Department of Education adjusted the 2013-14 assessment to reflect the new state standards.

- *Content* – The assessment was reduced from 67 to 60 questions based on a narrowing of the skills students were expected to demonstrate. Student Performance Indicators (SPIs) were dropped from prior TCAP tests to align with the new standards.
- *Proficiency Cut-Scores* – The percentage of questions that students needed to answer correctly to be considered proficient changed in grades 3 – 5.
    - In third grade, students needed to answer 75% of the questions correctly to be proficient in 2013, while 70% of the questions needed to be correct to be deemed proficient in 2014.

Below is a summary chart of the changes of item content for the TCAP from 2013 to 2014:

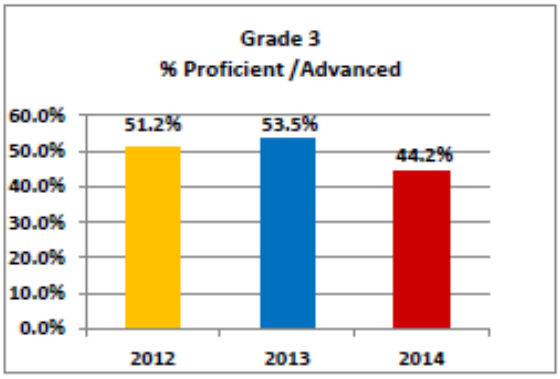
2014 minus 2013 Grade	Number of Questions	Language	Vocabulary	Writing and Research	Communication and Media	Logic	Informational Text	Literature
3	67 to 60	4.5%	1.0%	2.3%	1.2%	-0.4%	-1.8%	-6.8%
4	67 to 60	5.4%	-0.6%	-5.4%	-0.6%	-0.1%	7.7%	-6.4%
5	67 to 60	6.9%	-0.4%	3.0%	-2.1%	-0.1%	-2.3%	-4.9%

The KCS experienced a greater number of students who did not complete one or both of the reading sections. The district "omit rate" for the last four questions in each reading/language arts section increased significantly over prior years. As a result of similar trends across the state and educator feedback about the time needed for students to implement learning strategies they were taught for close reading of text, the 2015 TCAP will extend the time for reading/language arts by 10 percent.

As such, it is difficult to draw meaningful conclusions regarding achievement trends because the assessment will have changed over the three years comprising the early literacy program implementation.

However, achievement is distinct from growth, as growth is a relative measure over time. The TVAAS model does still provide valid prior year comparisons. The growth data for third grade is generated from the prior year performance on the SAT-10, and, therefore, the third grade TVAAS calculations use the same model as that of grades 1 and 2. The TVAAS measures for grades 1 – 3 are generated using a growth standard based on the *current year* statewide performance.

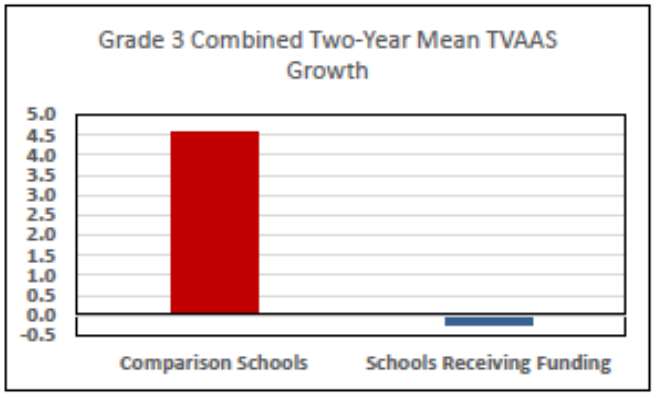
- **Achievement** * – Achievement for the district in grade 3 based on the percent of students who are proficient or advanced on TCAP increased over the prior year in 2013 and declined in 2014.



- **Growth** – There was positive growth in grade 3 for the district in 2013 and 2014. Combining results over two years, student growth was as predicted based on the TVAAS model for schools receiving early literacy funds.

	2013 One-Year TVAAS Growth	2014 One-Year TVAAS Growth
Subject	Grade 3	Grade 3
Reading/Language Arts	3.3	0.3

- **Growth Versus Comparison Schools** – The schools directly receiving supplemental early literacy funds in both years were outperformed by those schools with no funding in Grade 3. The schools are the same as those detailed in the section above, with the exception of the primary schools and the addition of Chilhowee Elementary.



* It is difficult to draw meaningful conclusions regarding achievement trends because the assessment will have changed over the three years comprising the early literacy program implementation.

Grades 4 and 5

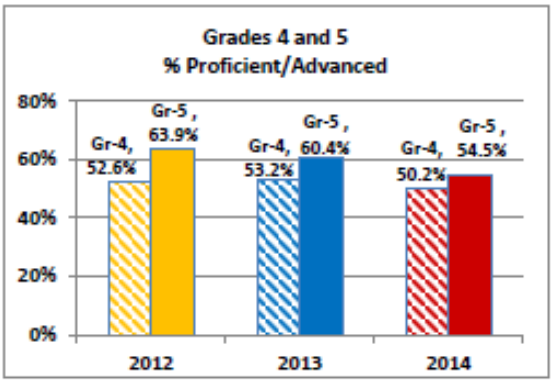
The fourth and fifth grade results are based on the TCAP reading assessment. As with the third grade assessment, the 2013-14 assessment was adjusted to reflect the new state standards. It is difficult to draw conclusions regarding achievement trends in fourth and fifth grade because the TCAP assessment has changed over the course of the last three years.

However, achievement is distinct from growth, as growth is a relative measure over time. The TVAAS model does still allow for valid prior year comparisons. The TVAAS measures for grades 4 and 5 are generated using a growth standard based on 2009 state-wide performance.

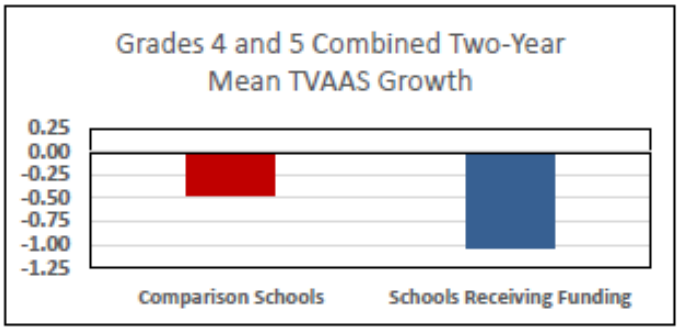
- **Achievement *–** In 2013, achievement for the district in grade 4 increased over the prior year based on the percent of students who are proficient or advanced on TCAP and declined in grade 5. In 2014, achievement in both grades 4 and 5 declined versus the prior year.

- **Growth –** Grades 4 and 5 had positive growth in 2013 for the district but growth declined in 2014 based on the TVAAS model. When reviewing the combined results over two years for schools receiving early literacy funds, students did not perform as well as predicted.

- **Growth Versus Comparison Schools –** The schools directly receiving supplemental early literacy funds in both years performed similarly to those schools that did not. The schools are the same as those detailed above, with the exception of the primary schools and the addition of Chilhowee Elementary.



	2013 One-Year TVAAS Growth		2014 One-Year TVAAS Growth	
Subject	Grade 4	Grade 5	Grade 4	Grade 5
Reading/Language Arts	2.1	0.0	-1.2	-4.1



* It is difficult to draw meaningful conclusions regarding achievement trends because the assessment will have changed over the three years comprising the early literacy program implementation.

Investment Analysis

The $2.87 million of additional funding was primarily allocated to teacher support personnel, including instructional coaches and instructional assistants, intervention and instructional materials, and professional development stipends and associated professional development costs such as pay for substitute teachers during Professional Development activities. (See Appendix B for *School Costs*.) The summary of these expenditures are below:

Early Literacy Initiative Expenditures	FY13	FY14
Instructional Coaches	$ 2,195,408	$ 1,801,202
Instructional Assistants	$ 316,785	$ 397,348
Other Personnel	$ 110,401	$ 115,914
Professional Development (Subs & Stipends)	$ -	$ 68,419
Intervention Materials	$ 200,000	$ 277,910
Other Materials	$ 44,904	$ 118,836
Carry-over (Available for FY15)	$ -	$ 90,371
Total	**$ 2,867,498**	**$ 2,870,000**

- **School Funding** – In FY13, 38 elementary schools benefited from this supplemental funding to support instructional coaching and instructional assistants. In FY14, 32 schools received the same. In FY13 and FY14, 27 and 21 schools, respectively, received more than $50,000 of supplemental funding from this effort.
- **System-wide Funding** – All of the district's elementary schools benefited from support of the Elementary Reading Supervisor, literacy trainers, and intervention materials for Voyager and Spire over the two-year timeframe.
- **Grade Level Emphasis** – The majority of funding for instructional coaching was focused in the early grades, which included more intensive coaching support for 15 schools. The additional instructional assistants were used to support students in grades 3 – 5 in 20 schools.

Continuous Improvement

The 2014 Educational Return on Investment (EROI) Report provided a detailed program evaluation for all the elements of the early literacy program, including investment analysis, findings and recommendations. The concluding section of this management report discusses efforts made to improve results in 2014-15 based on recommendations in the EROI report. Given the promising outcomes for grades K – 2 and the challenging results in grades 3 – 5, the instructional and school supervisors in curriculum and instruction have taken definitive action to accelerate results.

Early Literacy Recommendations and Action Steps

1) *Identifying and Prioritizing Students Targeted for Invention* – The Educational Return on Investment (EROI) evaluation determined there were many students placed in Voyager whose performance was well beyond the program's intended audience. With the implementation of the Response to Instruction and Intervention (RTI²) guidelines, the district has taken several steps to better identify students who need intervention. *Actions Taken:*

a. The KCS has implemented a new universal screener to replace AimsWeb CBM. This new STAR Renaissance assessment is a next-generation computer-adaptive assessment to help diagnose specific skill deficits for students.
b. All school teams were trained on how to make individual decisions on student placement in intervention. The district has a dedicated program manager to support RTI² implementation and coordinate the work of eight RTI² coaches to provide ongoing training and support to schools in the identification and monitoring process.
c. Schools are being discouraged from practices that included placing all third grade students in an intervention program as a response to the statutory requirements regarding third grade retention. Our school leaders better understand the potential harm of this practice, though well-intentioned. Intervention decisions must be specific to the individual student.

2) *Intervention Data Collection – EROI Recommendation*: The district should explore opportunities to record intervention data in student information systems (SIS) and/or our district learning management system. *Action taken:*
   a. The district has moved all progress monitoring to electronic format in STAR Renaissance (Tier 2) or AimsWeb 2.0 (Tier 3). Moreover, student enrollment in intervention is now being captured in the new ASPEN (SIS) system.
3) *Intervention Scheduling – EROI Recommendation*: The district should develop and offer supports to principals around optimal scheduling scenarios. *Actions taken:*
   a. The district provided administrators sample schedules created by three of their peers with the support of the Parthenon Group. Scheduling continues to be a concern, in terms of having enough personnel to support students not in intervention during that time.
4) *Intervention Programs and Fidelity Monitoring – EROI Recommendation*: The district should consider investigating other invention programs, as well as developing structures to monitor the fidelity of implementation of our intervention services. *Actions taken:*
   a. The Elementary Reading Supervisor established a taskforce of teachers and coaches to review intervention programs. The group produced a review of more than a dozen research-based intervention programs for Tier 2 and Tier 3 students in March 2014. This review now functions as the approved list of interventions from which schools may select materials.
   b. In addition to the comprehensive review of intervention programs, the district established a specific set of interventions for which materials would be purchased and district personnel would be trained to support. In addition to Voyager, which is now used exclusively for Tier 2 students, the district provides materials for *SPIRE-Sound Sensible, My Sidewalks,* and *Reading Street RTI Kits. Wilson Fundations* is also used on a limited basis where there are personnel certified by Wilson to deliver the intervention.
   c. As noted earlier, the RTI² implementation has established school-based committees specifically charged with fidelity monitoring. Bi-weekly and monthly routines are followed to evaluate student progress and program efficacy. There is also a district-wide team, including teachers, principals, coaches, and central office staff, which meets regularly to address any systemic issues or concerns.

First Grade Intervention Recommendations and Action Steps

1) *Quality of Classroom Instruction – EROI Recommendation*: Our program evaluation limitations acknowledged that the results for students in intervention are more heavily affected by the core instruction they receive from their classroom teacher. *Actions taken:*
    a. RTI² is designed to focus on instruction to bolster student results. The STAR Renaissance universal screener provides detailed instructional planning support for teachers, as well as recommendations for daily small group instruction.
    b. Increasing the access of schools to instructional coaching support is also a key component of this strategy. In schools where the first grade coach had a small team of teachers of 10 or less, the student growth significantly improved. We have made efforts to increase the availability of a full-time coach at all schools, as the budget has allowed.
    c. The district has established a *Year-Long Elementary Reading Course* that includes 11 half-day sessions. There are 10 cohorts of teachers from 19 elementary schools enrolled in the course. The KCS has two Reading Intervention coaches that have been trained in the "Teaching Reading is Rocket Science" workshops developed Dr. Louisa Moats. The course focuses on the science of teaching reading, understanding the progression of reading foundational skills and involves ongoing "bridge to practice" activities to connect the learning to daily classroom instruction. It is our intention that all teachers in grades K – 5 will be able to participate in this training to improve our collective expertise in reading instruction.

Additional Elem. Reading Support (Instructional Assistants) Recommendations and Action Steps

1) *Third Grade Results – EROI Recommendation*: Further qualitative investigation at individual schools should be pursued to ascertain why the results are so different (and disappointing) at the third grade level. *Actions taken:*
    a. In following-up with principals and coaches, it was discovered that many high needs schools were placing all third grade students in intervention in order to ensure that there would be no issues in terms of the retention requirements in state statute. Unfortunately, this practice did not help those students who did not actually meet the criteria for targeted intervention. In some cases, it even hurt those students' academic progress.
    b. As such, all principals received training that shared these specific findings and were discouraged from making blanket placement in intervention, as noted earlier.
2) *Quality of Intervention Personnel – EROI Recommendation*: Both teachers and principals indicated that instructional assistants were not as effective in delivering intervention supports. The district should consider if it is wise to continue to rely so heavily on instructional assistants to provide intervention services to students who are struggling the most. *Actions taken:*
    a. Principals received district training that highlighted this issue. As a result, the recommendation was made that only certified teachers provide Tier 3 intervention services. Instructional Assistants would only be involved in Tier 2, Voyager intervention, if possible. In most schools, we are able to ensure that Tier 3 interventions, for students who struggle the most, are delivered by classroom teachers exclusively. However, there are some high needs schools where this is not possible given the large number of students in Tier 3 intervention.

b. Many schools have begun to have instructional coaches work with Instructional Assistants (IAs) to bolster their skills in delivering interventions. When scheduling allows, IAs may participate in Professional Learning Communities to better monitor student progress. The district has also provided additional Voyager training to improve understanding and implementation of this Tier 2 intervention.

Conclusion

The early literacy program has been an important and beneficial investment in the reading skills of our youngest students. We continue to adjust, refine and improve this critical initiative based on program evaluation and student results. Leading indicators, especially in the primary grades demonstrate that this focused early literacy effort is having a positive impact on ensuring a strong academic foundation that will support the success of our students throughout their Knox County Schools experience, and throughout their lives. The promising results outlined in this report suggest that continued funding of the early literacy initiative by Mayor Burchett and the Knox County Commission will lead to better academic outcomes for our students, and therefore is a wise and necessary investment.