EDUCATION DATA COLLABORATIVE

Effective and Efficient: Aligning Data Systems and Research to Improve School Performance Research Report

Authors: Dr. Jimmy Byrd, Dr. Willard Daggett, Douglas Silver, Dr. Cynthia Williams
Project Director: Sylvia L. McMullen

March 2011

EDUCATION DATA COLLABORATIVE

PLANO – DENTON – ROUND ROCK – COLLEGE STATION – NORTH ZULCH

March 12, 2011

Dear friends and servants of our children and youth –

Too many of our children don't graduate high school; too many of our high school graduates are unprepared for college and/or the modern technological workforce; and too many of our aspiring scholars never emerge from college level "remedial" classes. The sentence precedent describes a lethal threat to our economic-competitive prosperity in Texas, the nation and the world.

Since serving on the Governor's Select Committee on Public Education in 1983-84 (HB 72) and through several other public "assignments" since, it has been my fond hope that we would be able to improve public education with a determined and palpable sense of purpose. Instead, in many instances and places we have witnessed a growing gap between what is expected of our children and what our globally competitive environment demands.

I am encouraged by recent initiatives to harness the capabilities of effective technology to support our schools' data driven decision-making. This in turn encourages accountability for improving performance. Through HB-SB 3, the 80th Session of the Texas Legislature tasked The Comptroller's office with overseeing the state's public and open enrollment charter schools education data management. Further focus was directed as same pertained to academic performance and financial efficiency. In partnership with the Education Research Center housed at The University of Texas at Dallas, the Comptroller's office recently released FAST as a first step in making accountability transparent to all stakeholders.

Over the last year, a group of committed educators has come together to design and implement a support system for school districts that not only assists in their accountability efforts, but also gives substantial support for their transforming accountability into responsibility. We believe the preceding, inevitably produces higher performance.

As a team, we sought to ensure that our Proof of Concept system would help schools identify practices, programs, and cost centers that are proven to be inefficient. We then recommended evidence based successful practices with which to make them more effective and efficient and partnered the Education Data Collaborative - a 501(c)3 entity - with DELL Services to address future scalability.

We feel quite strongly that determining effective and efficient rankings requires beginning within a school or district and understanding "why" and "what can we do inside" before we begin to compare "districts to other districts." Furthermore - and of greatest importance – we are committed to helping define and describe pathways for schools or districts to ascend in rank and provide tools to accomplish just that.

This report is only a beginning. As you read it, you will see that it gives credible hope that in the management and use of our students' "data" we can help significantly as they work to bring their dreams to life. In so doing, we will be helping to ensure our collective and bright future as Texans, Americans and world citizens.

We are indebted to our partner school districts, El Paso ISD and Socorro ISD and El Paso Community College, because without them this would not have happened. We are also eager to respond to your suggestions and answer questions as together, we continue to refine this effort. Welcome to this conversation!

Jon H. Fleming
Chairman and President

Acknowledgements

Special thanks to El Paso ISD, Socorro ISD, and El Paso Community College for their willingness and support to participate in this proof of concept project.

Additional thanks to the organizations who participated in the project: Data Driven Software Corporation (D2SC), Avid Stats, The Successful Practices Network, SchoolDude®, and to Krystal Drennan and Anna Walden for their invaluable technical assistance.

And…finally to our "partner" in this grand venture: DELL Services, Inc. They helped beyond telling in their consummate understanding of "things large"… which after all is the future of all this in Texas and beyond.

Table of Contents

Table of Figures

Table of Tables

INTRODUCTION

Preparing students for the 21st century workforce has been the challenge of the nation's education system for the last decade. Making data driven decisions has been the push for both districts and state departments. National emphasis has continued to be focused on high drop-out rates, low student performance, lack of rigor, and the need for reforming education. Most reports analyzed and stressed the need for changing high schools. In 2005, ACT focused the nation's attention on middle school education.[1] They noted that addressing the crisis in high schools begins with a student's experience in middle school. Studies have shown the link between Algebra 1 and high school success, the 9th grade "bubble" of failures, and the high cost of intervention at the high school level. Advanced data systems began to reveal the scope o the problems and were also seen as a critical part of the education reform.

> *"We will need to improve student performance if our students as individuals, and we as a nation, are to compete successfully in the 21st century global economy. Financially, schools will receive some temporary assistance with the federal stimulus package, but this will provide only modest and temporary relief. Therefore, schools must take the responsibility to find new and innovative ways to improve student performance with increasingly fewer resources."*
>
> *– Dr. Willard Daggett*

When the economic downturn began in 2008, data systems were also seen as a way to help schools cope with the financial crisis. In Texas, millions of dollars were spent trying to link, match, and merge data. Educators and legislators called for aligned data systems to facilitate analysis on college readiness, teacher effectiveness, on-track/off-track indicators, alignment of curriculum, and professional development. The Texas Education Agency refocused attention on improving school accounting systems and financial reporting. School accountability measures became the dominant focus of the education debate.

In response to this maelstrom, the Education Data Collaborative (EDC) evaluated these forces and designed a system of responsibility as an alternative. The operating theory[2] was the work of Dr. Willard Daggett[3] on efficiency and effectiveness released just after the nation plunged into a fiscal crisis, and the implementation of the federal American Recovery and Reinvestment Act (ARRA) stimulus package for education was beginning.

---

[1] Wimberly, George L. and R. Noeth. (2005) *College Readiness Begins in Middle School. ACT Policy Report. Retrieved from* www.act.org/research/policymakers/pdf/CollegeReadiness.pdf

[2] Dr. Willard Daggett is CEO of the International Center for Leadership in Education and Chairman of the Board for the Successful Practices Network.

[3] The full paper is attached as an appendix to the full EDC report and can be accessed at: http://www.leadered.com/whitepapers.html

This theory was based on two simple, yet conflicting challenges:

1. Improve student performance.
2. Deal with increasingly fewer financial resources than in the past.

These two challenges are the foundation for the Efficient and Effective Framework (Figure 1). The vertical line denotes cost of initiatives—or efficiency. The horizontal line represents student performance—or effectiveness—of an initiative. In the framework:

Quadrant A – represents low efficiency and low effectiveness

Quadrant B – represents high efficiency and low effectiveness

Quadrant C – represents low efficiency and high effectiveness

Quadrant D – represents high efficiency and high effectiveness

Initiatives in Quadrant D should be considered and those in Quadrant A should be questioned.

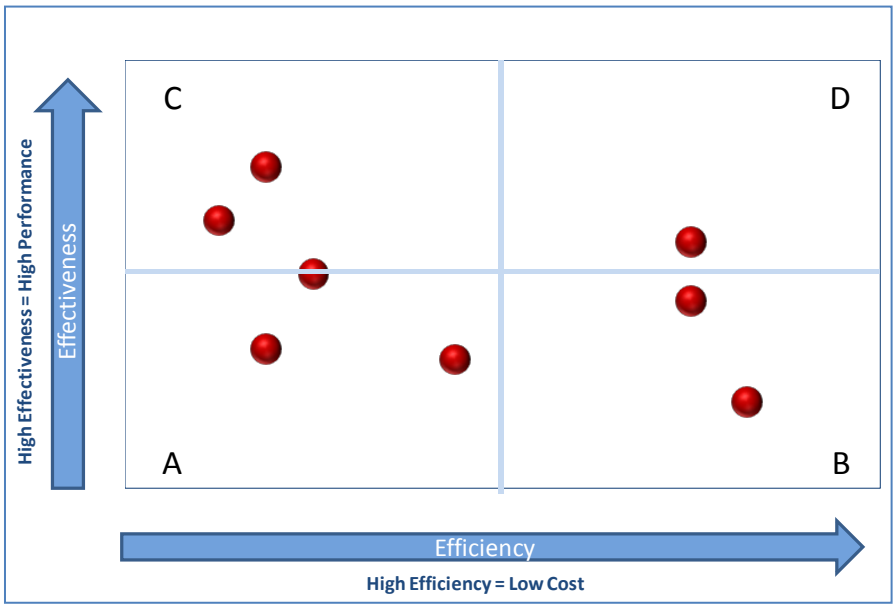

FIGURE 1. EFFICIENT AND EFFECTIVE FRAMEWORK

With the challenges facing schools and districts to do more with fewer resources, we sought to understand where efficiencies could be created at the middle school and classroom levels, clarify the link between student performance and resources, and seek potential areas of savings for schools and districts. Our goal was not to "fix" middle school education. Rather, we wanted to provide an alternative to the current model of "policy drives practice" and determine if "practice could drive policy."

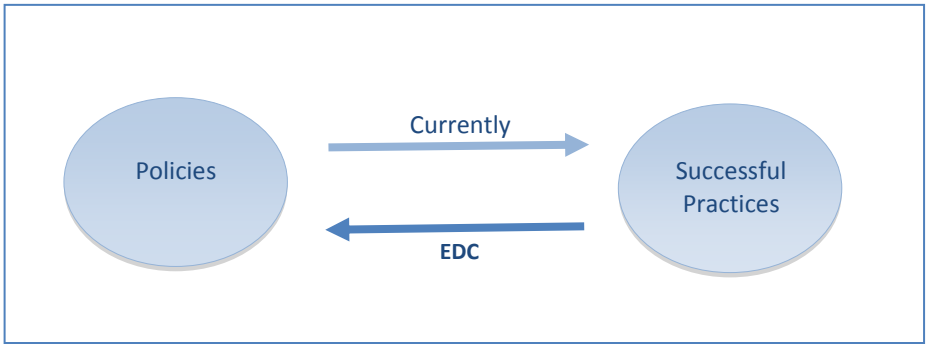

FIGURE 2. POLICIES AND PRACTICES MODEL ALTERNATIVE

Unfortunately, most existing data systems only supported focusing on state-level aligned assessment data for the purpose of accountability. Few schools and districts had implemented data systems designed as tools for classroom instruction, curriculum, and strong professional development in a real or near real-time data environment. Far too many schools (if not most) experienced "data lag" that could only target learning deficits missed weeks or months prior when the assessment was given, thus requiring typically more expensive and less effective intervention than timely prevention. *Collecting and analyzing state-level accountability data is important as an audit tool, but it has not helped identify pathways for schools to be both effective and efficient.*

It is our hope that the work of the EDC will provide optimism for schools and districts that what many have believed was not within their grasp is truly achievable. A single source of longitudinal data with user-friendly analytics of relevant real-time information is not a theory; it is a fact. Using data for improving educational performance and efficiency at the same time also is both the responsibility of the state and the district.

--Respectively submitted, Sylvia McMullen, Project Director

Executive Summary

Context

Texas Public Education is facing one of the most challenging academic and economic times in history. Texas has a graduation rate of 72.5%, slightly lower than the national average of 73.2%. College entrance exam scores have improved very little. Reading scores on the National Assessment of Educational Progress (NAPE) have remained static over the past two years, and schools have seen small improvement in math and reading TAKS scores. Statewide, there is a budget shortfall in Texas, which could impact over 15 billion dollars of education funding over the next two years. Estimates are that 100,000 teaching and administrative support positions will be lost before 2013. Public education accounts for approximately 44 percent of the state budget in Texas and over 80% of these funds are allocated to employee salary and benefits. Districts have indicated reductions will likely include cutting custodians and maintenance crew jobs, eliminating safety officers, reducing special education, cutting central office and technology personnel, reducing extracurricular activities, implementing tiered scheduling, closing some schools for consolidation within districts, and eliminating and consolidating some teaching and administrative positions.

The current accountability and data system in Texas focuses on lagging year-end data to evaluate student and teacher performance and does not provide teachers, principals, administrators, parents, students, and school boards with the ability to make real-time decisions that would allow for efficient and effective operation, or monitoring student growth within a year.

Few districts monitor student growth to determine points of intervention or to identify whether a student is on-track or off-track to transition to the next grade level, to graduate high school, and graduate post-secondary ready. Further, the current system has limited capability to determine whether or not schools are operating as efficiently and effectively as possible by using only state released data rather than data from source systems. Figure 3 below is an example of an on-/off-track report for an Algebra 1 class, tracking students for college and career readiness throughout the school year. The algorithm can also be used to measure teacher effectiveness.

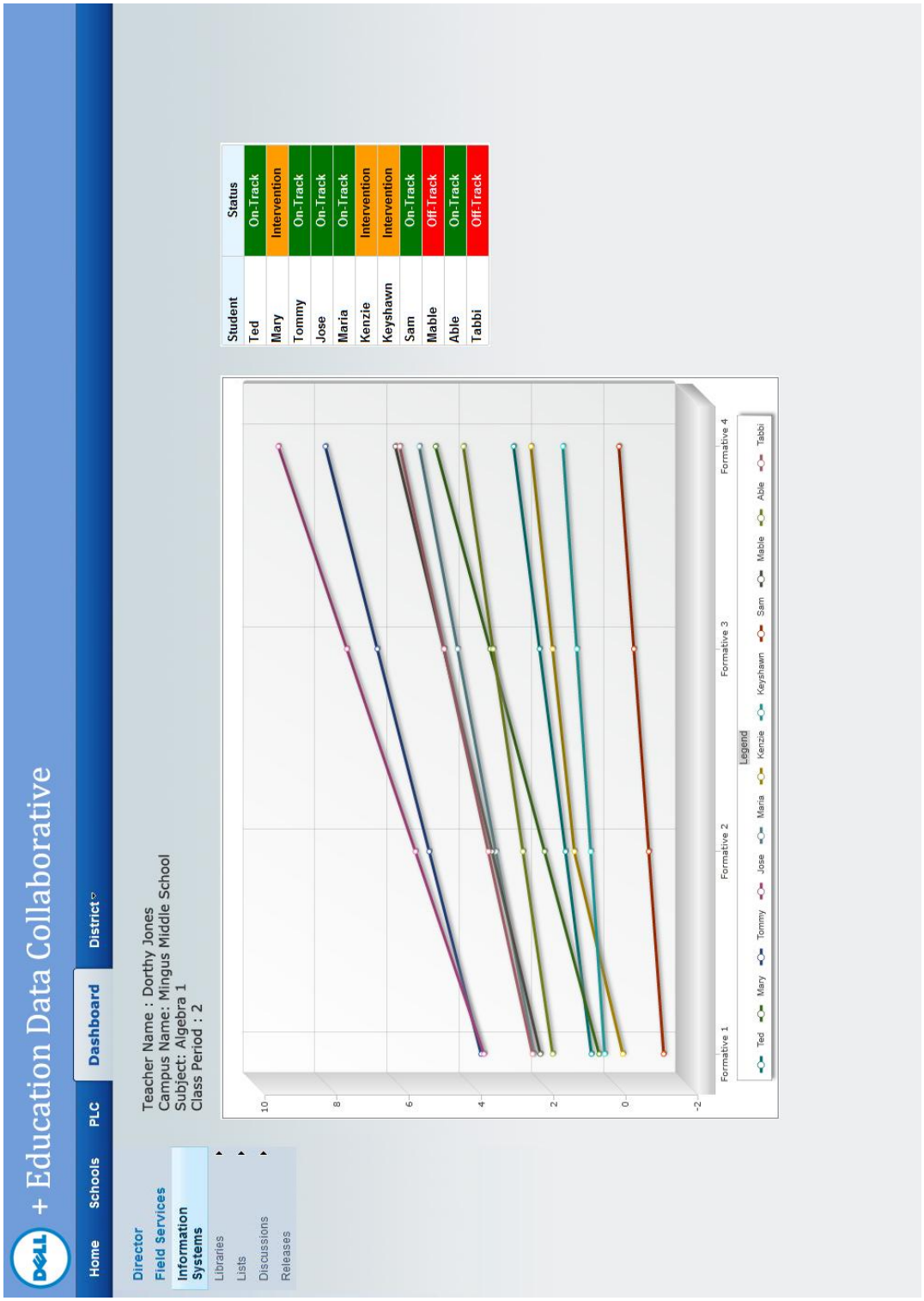

FIGURE 3. ON-/OFF-TRACK REPORT

The budget crisis is focusing the discussion on efficiencies and the need for development of accountability systems from real or near real-time data source systems at the campus and district level.

Rationale and Purpose

Districts are under immense pressure to efficiently cut budgets, but few have the ability to analyze district data and validate whether or not the budget cuts they make will impact performance outcomes. Currently data is available at the state level which provides aggregate numbers and, thus, only allows districts to compare one district to another district. The proof of concept pilot study associated with this project was designed to think **differently** about the challenges facing districts and to assist schools and districts with data from source systems in a real or near real-time basis that can be utilized to make within district comparisons and decisions. This proof of concept pilot study was designed to assess the effectiveness and efficiency of schools at the classroom, teacher, campus, and district level.

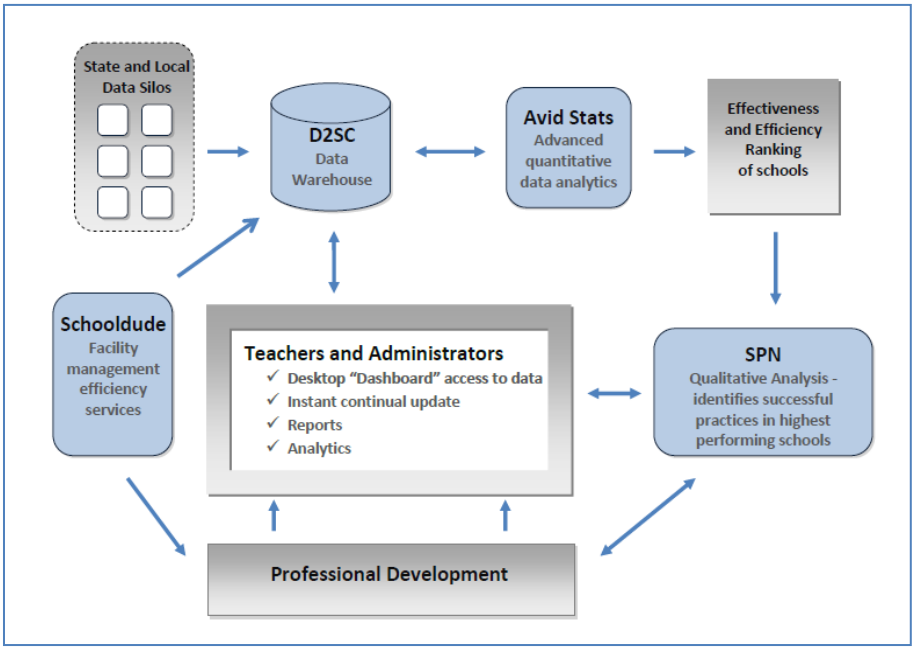

Figure 4. EDC Process Flow

This model becomes sustainable when — as the proof of concept did — a single database of all available data creates a longitudinal student record. The resulting real or near real-time data allows for monitoring of student, teacher, and school performance to determine the moment when a school transitions from effective to ineffective. By identifying when a student is "on-track" or "off-track", *immediate preventive steps can be taken*. This real time, daily, data monitoring provides decision makers with powerful, actionable information.

The focus of this work is on efficient prevention rather than expensive intervention in assisting schools' existing efforts and identifying potential areas of savings. These analyses and subsequent reforms become the successful practices that others can emulate to realize similar results.

The focus of this work is on efficient prevention rather than expensive intervention in assisting schools' existing efforts and identifying potential areas of savings. These analyses and subsequent reforms become the successful practices that others can emulate to realize similar results.

Texas' need for a longitudinal data system is well-documented by the legislature, the Texas Education Agency, the Comptroller's office, the Texas Data Collection, Analysis, and Reporting Systems Investigation (TDCARSI) Report, the longitudinal grants from the United States Department of Education, and the proposed development of the Texas State Data System. The proposal for the District Connections Database required the extraction of data from the source systems for both accountability reporting and providing actionable data for classroom teachers within the districts.

The willingness and support of El Paso ISD (EPISD), Socorro ISD, and El Paso Community College to participate in this pilot study was testament to the need for real-time data and the commitment to operating efficiently and effectively to improve the outcomes of all students in Texas. The proof of concept created not only a single database of all available data, it also created a longitudinal student record. The resulting real or near real-time data can allow for the monitoring of student, teacher, and school performance to determine the moment when a school transitions from effective to ineffective, as well as when a student is on-track or off-track and in need of immediate *preventive* intervention.

The purpose of the project was three-fold:

1. To provide near real-time data to classroom students, parents, teachers, principals, and administrators throughout the school year regarding student growth and progress. This information was expected to allow teachers to make formative, data-driven decisions regarding student instructional needs. Further, it was anticipated that teachers would be able to monitor all students in a real or near real-time basis for on-track/off-track and post-secondary readiness.
2. To provide near real-time data to principals that can be utilized to monitor and ensure that teacher effectiveness is linked to student performance in a fair, equitable, and transparent method.
3. To identify areas of efficiency and effectiveness at the campus and district level. Through quantifiable and qualifiable data validation, this information will allow districts to identify campus-level modifications which can improve all performance.. This study was designed to begin at the grass-roots level and work within districts to then help improve accountability between districts.

Literature Review

In response to No Child Left Behind, an extensive body of literature now exists that identifies characteristics and practices of effective principals and teachers. More recently, research on the efficiency of schools and districts has begun to emerge as well. This literature review outlines current research in these three areas: effective teaching practices, effective principal practices, and efficient school/district practices.

Effective Teachers

With regards to teacher effectiveness, the works of Hattie (2009), Marzano (2007), Marzano, Pickering and Pollock (2004), and Stronge (2007) were compared and contrasted. In total, these researchers examined over 55,000 pieces of evidence as related to teacher effectiveness. Strong (2007) identified 27 characteristics of effective teachers. Marzano et al. (2004) identified 9 practices of effective teachers. Marzano (2007) identified 10 practices of effective teachers through analysis of the meta-analyses of Haycock (1998); Marzano (2003); and Nye, Konstantopoulos and Hodges (2004). Lastly, Hattie (2009) synthesized over 800 meta-analyses relating to high levels of student achievement.

Collectively, the aforementioned research revealed nine areas of effective teaching practices which are believed to be directly linked to positive student outcomes. Common practices and characteristics of effective teachers include: they have passion to educate all students, encourage students to give their best effort each day, and personally strive to leave no child behind (Hattie, 2009; Marzano, 2007; Marzano et al., 2004; Stronge, 2007). These teachers believe that failure is not an option and utilize an extensive instructional strategy bank on a day-to-day and student-to-student basis. Effective teachers are proactive about understanding district curriculum, utilizing data management systems, and seeking professional development opportunities which can help extend their instructional skills. Further, these teachers understand that each student learns in a unique way, and make certain every child knows he or she is genuinely cared for. Although highly qualified teachers understand No Child Left Behind mandates that all children have the right to learn, effective teachers ensure that all children have the tools they need to succeed (Hattie, 2009; Marzano, 2007; Marzano et al., 2004; Stronge, 2007).

Effective Principals

Characteristics of effective principals are similar, though the umbrella of principal responsibilities spans beyond classroom walls. To examine school leadership practices, Marzano, Waters, & McNulty (2005) conducted a meta-analysis of 69 studies in which specific proficiencies related to the importance of principal leadership were explored. As a result of this meta-analysis, researchers isolated 21 responsibilities of a school principal which were positively linked with increasing student achievement (Marzano et al., 2005). Of the 21 responsibilities identified, ten were determined to be of particular importance and are outlined below.

Effective principals adapt to change and create reasonable school routines for faculty, staff and students. They encourage and offer teachers meaningful professional development opportunities, and strive to create a vision and positive educational culture which is shared by all. These school leaders evaluate teachers with college readiness in mind, and unite campus stakeholders on college readiness initiatives. Effective principals implement change which builds relationships by

creating and monitoring behavioral systems that do not interfere with student achievement. Effective principals stay abreast about district curriculum, assessment, accountability, policy, district data management systems, and instructional strategies (Marzano et al., 2005).

Efficient School Districts

Building an efficient school community where faculty, staff, and administration maintain strong partnerships and support a common vision for student success is a multifaceted challenge (Leithwood, Harris, & Strauss, 2010; Zimmerman, 1991; Calkins, Guenther, Belfiore, & Lash, 2007; Tough, 2009). Achieving the aforementioned in a cost-effective manner is yet another challenge. Currently this appears to be a weaker link in the literature, as studies regarding district efficiency are not as prevalent as the aforementioned research regarding teacher and principal effectiveness. However, research does suggest there are strategies which efficient, effective districts consistently utilize.

Efficient districts and campuses are proactive at the outset. They recruit a large candidate pool to include minorities, complete stage-wide reductions of applicants using various pieces of evidence, and promote their districts to the most qualified applicants (Darling-Hammond & Berry, 2006; Peterson, 2002). Beyond hiring, efficient districts strive to retain teachers by offering incentives such as tuition assistance, child care assistance, and better pay (Johnson, 2000; Temin, 2002); and they seek teacher input when making district-wide decisions (Goldberg & Proctor, 2000). Further, because districts realize that new teachers and principals are generally placed in the most demanding classrooms due to seniority policies, efficient districts implement proven mentoring programs to support new educators in response (Guarino et al., 2006; Darling-Hammond & Berry, 2006; Ingersoll, 2001).

Efficient districts strive to establish a readiness to learn, a readiness to teach, and a readiness to act in all campuses by providing funding, mentoring, and professional resources needed to support campuses, especially those in need of turnaround (Calkins et al., 2007). Effective districts strive to establish positive academic cultures in all campuses and among all school stakeholders. Further, in response to underperforming campuses, effective and efficient school districts proactively work to turnaround these schools, and realize the transformation from ineffective to effective is an arduous task which can take significant time (Tough, 2009; U.S. Dept of Education, 2004; Calkins et al., 2007; Rivero, 2009).

With regards to accountability, efficient districts view assessment as a useful tool that helps teachers, principals, and district administrators identify areas of strength as well as areas which need strengthening (Tough, 2009). Efficient districts embrace formative and summative evaluation as assets in a variety of situations, including organizing instruction, developing curriculum, managing school day-to-day functions, strategic planning for the future, and developing behavioral plans (Tough, 2009). Efficient districts maintain central data management systems which include on- and off-track early warning systems, accurate student-teacher information, and links to post-secondary success.

In conclusion, the review of literature indicates that in order for teachers and principals to reach their full potentials as effective educators, they must position themselves in efficient schools and supportive districts where reciprocal relationships between teachers, principals, staff and district administration are encouraged and valued. Successful schools and districts embrace climates of

collaboration, possess a willingness to make to change as needed, exemplify a spirit of fairness and respect for all, and exhibit a can-do attitude which motivates teachers, staff, principals, and district administration alike. When these elements exist, effective educators truly realize they are invaluable and empowered assets of successful school communities, allowing districts to operate most efficiently.

METHOD

A single database was generated to allow for quantitative statistical models and predictive analysis in three specific areas: readiness indicators, teacher effectiveness, and efficiency/effectiveness at the classroom, school, and district level. Further study was conducted to qualitatively confirm successful practices in high performing efficient and effective schools for replication in schools not performing at peak efficiency and effectiveness. Although results from the pilot study cannot be generalized beyond the EPISD, the model itself is scalable for any other district, region and/or state based on the same applicable processes.

This single database approach has additional, attractive features. First, this approach can provide for seamless student records in a secure, FERPA compliant environment. These records can be electronically exchanged between schools, districts, and community colleges to address student mobility factors and allow schools to quickly place students in appropriate classes. This approach can also provide a single longitudinal student record for the consistent educational relevance of every individual student. Further, the single database approach can include community college data, which can provide analytic feedback on student performance and curriculum. This feedback can serve to inform instructional change at K-12 levels with the goal of reducing the need for developmental or remedial classes at the post-secondary level.

DATA COLLECTION

The project was implemented in two quantitative phases, and one qualitative phase. Phase one focused on establishing a longitudinal database and analyzing data necessary to support the project. Phase two focused on implementing the campus, teacher, and student-level components of the study. Data related to students, staff, finance, and student achievement were examined. Relevant data and results were reported and shared with EPISD in a meaningful, easy-to-use, format that was intended to inform and drive daily instruction, staff development, and organizational improvement.

Data Driven Software Corporation (D2SC) was engaged to extract the data from multiple sources. The model/process was implemented in multiple modes including near-real-time and batch. The process facilitates extracting, mapping, and loading data from multiple source systems (SIS, HR, Finance, as well as external systems and resources). For the current EPISD implementation, the 'batch' mode was implemented and included the following EPISD-housed data source systems:

- TEAMS (SIS)
- CIMS (HR + Financial)
- Assessment [multiple systems and sources including: state mandated Texas Assessment of Knowledge and Skills (TAKS), District Benchmark System, etc.]

For various data housed in systems (CIMS, TEAMS, etc.), those systems were checked for updates regularly and/or as desired by EPISD. D2SC's Data Warehouse provided a single-site repository for all transactional data as well as various summary and calculated data.

Research Design

A mixed method research design was employed where quantitative analyses were initially conducted. Qualitative analyses followed, with the goal of qualitative analyses being to verify and provide context to quantitative outcomes. Further, qualitative results also served to inform the need for additional quantitative analyses. For the first two purposes and phases of this project, the unit of analysis was primarily at the student level, with non-random classroom assignment. For the third purpose and phase of this project, the units of analyses were primarily the campus and district levels, respectively.

Participants

EPISD is a public school district located near the Texas-New Mexico border and in the El Paso Community College (EPCC) vicinity where the majority of area high school graduates attend. EPISD is the 10th largest district in Texas and the 61st largest district in the United States. EPISD educates more than 63,000 students in 94 campuses, with more than 81% of the student body identified as Hispanic. With almost 9,000 employees, EPISD currently has an annual operating budget of $475 million.

Variables Considered in the Project

Research indicates the efficiency and effectiveness of any school district is multifaceted. In order gain a comprehensive understanding of EPISD, the following variables were selected for quantitative examination in conjunction with this study: academic outcome variables, student demographic variables, teacher variables, school variables, campus variables, and district variables.

Academic outcomes include variables such as grades and test scores. Student demographic indicators include variables such as socioeconomic status, at-risk or special education classifications, and ethnicity. Teacher indicators include variables such as type of certification, years of experience, and number of days on campus and teaching in a classroom. School indicators include variables such as campus mobility rate, demographic distribution of students, and student-teacher ratios. Campus finance indicators include variables such as expenditure per student as well as expenditures per instruction or intervention type. District finance indicators include variables such as district expenditure per pupil, district expenditure on school leadership per pupil, and district expenditure by instructional program per pupil. An exhaustive list of all 68 variables quantitatively examined can be found in Appendix B.

Quantitative Procedure

Effectiveness

Initially, EPISD provided data, which was input into the D2SC data warehouse, and extracted for analysis. Then, a longitudinal data file was constructed at the student level that included student demographic variables, six-week averages in each all courses that were linked to the teacher who taught the course, benchmark assessment results in reading/ELA and math from each of the three benchmark assessments administered within the school year, and math and reading TAKS scores, which included scores from the last three administrations of TAKS (primary administration). Descriptive measures of each of the data elements were

calculated to ensure there were no erroneous entries and to verify data integrity. Note that SAS 9.2 was utilized for all quantitative analyses.

The initial analyses focused on determining how well benchmarks were aligned with six-week averages in reading, mathematics, and TAKS. The rationale for this procedure was to determine if current benchmarks could be used to make near real-time adjustments in instruction during the school year and serve as a measure of teacher effectiveness.

After considerable analysis student results on benchmark assessments were standardized and linearly equated to allow the research team to use the measures in a growth model. Next, a multilevel growth model that included time at level-1 and student-level covariates at level-2 (See the variables list in Appendix B.) was generated for each campus to determine the amount of growth or gain among individual students in reading and mathematics. The base model employed in the study is below.

Growth Model (1)

$$Y_{ti} = p_{0i} + p_{1i}T_{ti} + e_{ti}$$
$$p_{0i} = b_{00} + u_{oi}$$
$$p_{1i} = b_{10} + u_{1i}$$

where,

Y_{ti} = outcome for subject i at time t.
p_{0i} = status of subject i when T =0.
p_{1i} = average change per unit of T for subject i.
b_{00} = grand mean of status when T =0.
b_{10} = grand mean change per unit of T.
e_{ti} = within person error.
u_{oi} = unique increment to grand mean of status for subject i.
u_{1i} = unique increment to grand mean of change for subject i.

The amount of growth or gain between each benchmark assessment was then entered into an algorithm to determine if a student was on- or off-track to pass TAKS in reading and mathematics during the school year. This measure provided points of intervention that were utilized to determine which students were on-track to pass TAKS and to assist teachers in providing immediate assistance to those who were identified as off-track to pass math and reading TAKS. This approach differed from current practice where campuses must wait until the end of the year to determine how students perform on these same assessments. Prior work using this algorithm was proven to be 94% successful in the state of Arkansas and in other districts that have beta tested the algorithm in Texas. With this information, a district can generate reports showing each students' on- or off-track progress, which can subsequently be shared with students, parents, teachers, and the administration.

Next, to model growth in student TAKS performance in math and reading, TAKS results from each student over a three-year period were assembled for both reading and math and examined. Due to the inherent psychometric problems associated with the TAKS, student results were standardized and linearly equated. Following this, teacher effectiveness scores were calculated for each reading and math teacher in each middle school in EPISD. Two models were calculated to compare teacher effectiveness scores:

- Model 1 was based on benchmark assessments in reading and math that were administered within the school year; and
- Model 2 was based on student performance on reading and math TAKS assessments which were administered and reported at the end of the school year.

The rationale for comparing the two models was to determine the utility of measuring teacher effectiveness within the school year . The rationale for modeling teacher effectiveness within the school year, based on near real-time data, was to assist principals in identifying areas of professional development needs during the school year, instead of after the school year, so that they would be better equipped to implement meaningful professional development in a timely manner. The models were estimated using maximum likelihood. The (unadjusted) teacher effects, r_{0jk}, were predicted based on estimated variance components and were reliability-weighted. These calculated effects were the empirical Bayes residuals, and formed the basis for estimating teacher effects. The base model utilized during this phase of the study was of the following form:

Level- 1 (Student Level) : $Y_{ijk} = \pi_{0jk} + \sum_{i=1}^{i} \pi_{jk} X + e_{ijk}$ (2)

where,

i = 1,…,n students
j = 1,…,n teacher
k =1,…,n campus
Y_{ijk} = student TAKS reading or math score
π_{0jk} = student-level intercept
π_{jk} = student-level coefficients
X= student-level control variables
e_{ijk} = student-level random error, with e_{ijk} ~N(0;σ2)

Level- 2 (Teacher Level) : $\pi_{0jk} = \beta_{00k} + r_{0jk}$, (3)

$\pi_{ljk} = \gamma_{l00}$ where l = 1……..n.

where,

β_{00k} = teacher-level coefficients
γ_{l00} = non-randomly varying intercepts
r_{0jk} = teacher-level random effect, with r0jk ~N(0; τ22) for teacher j nested in school k.

The level-3 model allows for the clustering of teachers within campuses:

Level- 3 (School Level) : $\beta_{00k} = \gamma_{000} + u_{00k}$ (4)

where,

γ_{000} = non-randomly varying intercept
u_{00k} = campus-level random effect, with r_{0jk} ~N(0; τ_{22}) for school k.

Efficiency

The second phase of the study focused on the efficiency of each campus within a district in achieving student academic growth in core subject areas. Data Envelopment (DEA) and Stochastic Frontier Analysis (SFA) were employed to provide a summary where dollars were spent and to identify areas for reducing expenditures. The goal of this task was to provide areas of potential savings while achieving maximum output or growth given current inputs or expenditures.

Initially, student-level data were obtained from EPISD for all middle schools examined in the study from the 2007-08, 2008-09, and 2009-10 academic years. Data obtained included student demographic information, six-week grades in all subject areas, results on benchmark assessments in reading and mathematics and TAKS results in reading, mathematics, writing, and science. In addition, complete budget and human resource data for each campus were included. Subsequently, all instructional data files (student demographic data, six-week grades, benchmark assessments, and TAKS scores) were linked together based on student and teacher identification numbers. The complete data file provided a longitudinal record of each student and teacher over a three year period. In addition, campus budget and human resource data were assembled to provide a three-year record of expenditures and human resource information for each teacher at each individual campus. Subsequently, the data were examined in-depth to ensure the records were complete and free of erroneous entries. Finally, the variables for the efficiency study were assembled for analysis.

DEA was conducted in two stages. In stage one, an input and an output oriented CCR model was calculated to obtain efficiency scores for each middle school that included budget expenditures (See Appendix C.) and the percentage of students passing all TAKS.

Then, utilizing Tobit regression, the efficiency scores obtained from the DEA analysis were regressed on selected factors not controlled by the schools that were hypothesized to have a significant impact on the calculated efficiency scores. McCarty et al. (1993) suggest using efficiencies generated by (DEA) as dependent variables in a second stage with Tobit regression to assess the effects of variables not included in the first stage on technical efficiency. Since the efficiency estimates from the first stage are between 0 and 1, data is censored, and therefore Tobit regression, rather than OLS, is the appropriate method of estimation. In order to obtain efficient parameter estimates, the possibility of the existence of heteroscedasticity in this stage was considered and incorporated into the model. Further, to contrast DEA, a Stochastic Frontier Analysis (SFA) model was calculated to obtain efficiency scores for each middle school. The DEA and SFA scores were entered into the researcher's database and compared. Finally, cost comparisons and areas for potential savings were examined. While these measures are not new, they have not been commonly applied to education. Therefore, a brief overview of the DEA and SFA is provided below.

Overview of Efficiency Models

The measure of efficiency is normally one of either output efficiency (the distance between observed and maximum possible output for given inputs) or input efficiency (the distance between observed and minimum possible input for given outputs). The method to implement this analysis was both stochastic and deterministic. The former allows random noise due to measurement errors, while

the latter attributes the distance between an inefficient observed school and the efficient frontier entirely to inefficiency. The most popular methods to measure efficiency include SFA), which is stochastic and parametric, and DEA, which is deterministic and non-parametric.

Data Envelopment Analysis

DEA is a widely applied linear programming-based technique with the primary purpose of evaluating the efficiency of a set of decision-making units. DEA has mostly been used for benchmarking and for performance evaluation purposes due to its numerous advantages. The advantages of DEA are that:

- it allows inputs and outputs to be expressed in different units of measurement.
- It does not require an assumption of a functional form relating inputs to outputs.
- Decision making units or schools are directly compared against a peer or combination of peers.
- Efficient units form the "efficient frontier" and inefficient units are enveloped by this frontier providing information on their improvement potential.

In the current study, an input- and output-oriented DEA model was calculated. Input-oriented DEA models define efficiency as "the least input for the same amount of output." In the current study, the research team was interested in determining how much money could be saved while achieving the same output, where output was measured by individual student growth on state-mandated assessments during the timeframe of the current study. With input-oriented DEA, the linear programming model is configured so as to determine how much the input use of a school could be reduced to efficiently achieve the same output level. The procedure of finding the best virtual producer can be formulated as a linear program. Analyzing the efficiency of *n* producers is then a set of *n* linear programming problems. The following formulation is one of the standard forms for DEA (Charnes, Cooper, & Rhodes, 1978).

$$(1)\ Max\left\{\theta_0^* = \frac{\sum_i \mu_i y_{i0}}{\sum_j v_j x_{j0}}\right\} \qquad (5)$$

$$Subject\ to:$$

$$(2)\ \frac{\sum_i \mu_i y_{ik}}{\sum_j v_j x_{jk}} \leq 1\ for\ all\ DMUs\ k=1,2,\ldots,n$$

$$(3)\quad \mu_i \geq 0$$

$$(4)\quad v_j \geq 0$$

Where θ_0 = the efficiency score of the DMU under analysis;
n = number of DMUs under analysis;
I = number of outputs;
J = number of inputs;

Y_k = { y_{1k}, y_{2k},...,y_{ik},...,y_{Ik} is the vector of outputs for DMU k with y_{ik} being the value of output i for DMU k;

X_k = { x_{1k}, x_{2k},...,x_{jk},...,y_{Ik} is the vector of inputs for DMU k with x_{ik} being the value of input j for DMU k;

μ and ν the vector on multipliers respectively set on Y_k and X_k where u_i, v_j = the respective weights for output i and for input j;

Given a set of J Decision Making Units (schools), the model determines for each DMU_0 the optimal set of input weights $\{v_{i0}\}_{i=1}$ and output weights $\{\mu_{r0}\}_{r=1}$ that maximizes its efficiency score e_o.

Output-Orientation

In the output-oriented model, efficiency is identified as a proportional increase in the output vector (student growth) with a given input vector (budgeted expenditures). Therefore, the output-oriented measure of efficiency is the solution to the following constant returns to scale (CRS) DEA linear model (Coelli et al., 1998):

$$\begin{aligned} &\max \phi, \\ &\phi,\lambda \\ &s.t.\ -\phi y_i + Y\lambda \geq 0 \\ &x_i - X\lambda \geq 0 \\ &\lambda \geq 0 \end{aligned} \qquad (6)$$

where ϕ is a scalar, and y_i and x_i are column vectors of outputs and inputs respectively for the ith school. λ is an N x 1 vector of constants. The variable Y is an M x N output matrix and X is a K x N input matrix, and the proportional increase in outputs that could be achieved by the ith campus, holding inputs constant, is ϕ-1($1 \leq \phi < \infty$) and $1/\phi$ is the campus' efficiency score which is between 0 and 1.

While DEA is a popular tool to analyze efficiency, there are disadvantages. DEA is intended for estimating the relative efficiency of a DMU or school but it does not specifically address absolute efficiency. In other words, it tells how well the DMU or school is doing compared to the *peers* (set of efficient units), but not compared to a theoretical maximum. In addition, DEA is an extreme point technique, noise such as measurement error can cause significant problems. Finally, DEA is a non-parametric technique, making statistical hypothesis test difficult.

Stochastic Frontier Analysis

In addition to calculating the input- and output-oriented DEA model, an SFA model was calculated to provide further insight and verification of the DEA analysis. This approach uses econometric techniques and imposes a priori the functional form for the frontier and the distribution of efficiency. To better understand the function of the SFA model, consider a production function for n DMUs and K inputs of the following form:

$$y_i = \alpha + \sum_{k=1}^{k=K} \beta_k x_{ik} + e_i \quad i = 1 \ldots n \qquad \text{K inputs} \qquad (7)$$

where y is output, x_{ik} are inputs, and e_i is the residual for DMU or school i.

It is the residual e_i that captures any inefficiency in this model, while also capturing other noise or random effects (e.g. omitted variables, measurement error). SFA attempts to decompose the error term into inefficiency and noise components for each DMU or school i as shown in equation 8.

$$\ln y_i = \alpha + \sum_{k=1}^{k=K} \beta_k x_{ik} + [v_i - u_i] \quad i = 1 \ldots n \tag{8}$$

In equation 2, the error term is decomposed into two components, namely v_i and u_i, where *v* is an identically distributed conventional two-sided error term with zero mean and measures ***random noise*** and *u* is an identically distributed one-sided error term with a non-zero mean and ***measures inefficiency***. Note *u* is typically assumed to be exponential, half-normal or truncated normal. Finally, technical efficiency (TE) is measured as

$$TE_i = \frac{y_i}{\exp(x_i\beta + v_i)} = \frac{\exp(x_i\beta + v_i - u_i)}{\exp(x_i\beta + v_i)} = \exp(-u_i) \tag{9}$$

Qualitative Procedure

Using the quantitative effective and efficiency rankings of the schools in this project, two of the highest performing schools and three of the lowest performing schools in EPISD were selected for qualitative examination. These schools were identified on the basis of related quantitative analyses as well as the recommendation of EPISD administration. The intent of this analysis was to verify quantitative analyses associated with this study, and to develop an understanding of the conditions and practices that contributed to the quantitative ranking of each school. Although data collection was consistent from school to school for many of the questions and data sources, some differences in qualitative data collection among site visits did exist.

Two to three person research teams conducted extensive fieldwork over the course of two days at each of the five selected schools. The size of the research team depended on the size of the school and areas of focus as determined by quantitative analyses. Data collection employed three techniques: document analysis, semi-structured interviews with a variety of campus staff and administration, and classroom observations. These techniques were used as a method of triangulation in an attempt to increase the trustworthiness of the study (Maxwell, 2005).

The interview process involved focus groups and individual interviews with separate staff groups. This technique allowed for both the observation of participants interaction on a topic as well as an in-depth understanding of a person's opinions and experiences (Morgan, 1997). The selection of staff focus group participants was done through random selection by the school. Researchers conducted 6-8 teacher focus groups per school as well as a minimum of two individual teacher

interviews. The leadership team at each school also participated in a semi-structured interview.

All focus group interviews followed a semi-structured interview protocol in order to add consistency to the process. The interview process involved the same two researchers through the entire research project. All interviews lasted a minimum of forty-five minutes. Interviews were all audio recorded, detailed notes were taken, and interviews were transcribed.

Researchers also conducted classroom observations. These observations sought to understand the types of instruction common in the school, the types of learning activities students were engaged in, and the dynamics of effective learning. Classroom observations lasted anywhere from ten to thirty minutes, depending on the lesson. The number of classroom observations per school was dependent on school schedules and availability, but typically about fifteen classes per school. Immediately following the completion of each school visit the researchers shared and discussed their ideas in order to assist the coding process. Documents were collected at each school site as appropriate to the findings derived from the quantitative study. Examples of the document types include: schedules, websites, handbooks, course brochures and descriptions, and certain TEAMS reports. In addition, teachers and counselors also gave the researchers material they felt were relevant to understanding their school. Following each school visit, researchers organized and submitted notes and summary findings.

Data Analysis

Data analysis was iterative throughout the process. Researchers continually read notes, discussed findings, and asked questions. Researchers input, transcripts, and documents were used to develop codes. The first level coding consisted of creating a matrix to ensure that all interview questions were answered. Specific missing data was requested from school leaders. When necessary, researchers asked for administrative or teacher input on questions that were not fully answered. The second level of coding was deductive and involved codes developed from the research questions. Specifically, codes were developed related to key themes (e.g., success factors, academic culture, leadership, and classroom rigor). A third level of coding was concept driven (Gibbs, 2007). The concepts originated from the quantitative analysis. Finally, pattern coding was used to pull together all the levels of coding. This project conducted a cross case analysis using a matrix to organize individual schools according to the cross case themes.

Trustworthiness

The researchers followed Creswell's (2003) strategies for ensuring the trustworthiness of this study. These strategies included triangulation; member check; rich, thick descriptions; clarification of researcher bias; peer review and debriefing; negative case analysis; and external audit. The only strategy not employed in this study was an external audit. The strategy of negative case analysis was significant for the trustworthiness of this study (Maxwell, 1994). When the researchers found a theme in one of the schools case different from the others, it was analyzed rather than dismissed.

TESTING + ACCOUNTABILITY = WASTED TIME

With the passage of No Child Left Behind (NCLB), high stakes assessments began to play a significant role in all public schools across the United States. Since that time, schools have continuously searched for progress monitoring options to identify students at risk of not meeting a minimum goal prior to the state assessment. These options include the use of benchmark assessments. Often schools construct and use benchmark assessments to measure student progress with little or no data to determine the utility of the benchmark tests. Of particular interest in the current study was the utility of benchmark assessments to provide quality information in monitoring student progress. The results from the quantitative analyses of EPISD benchmarks in middle school reading and mathematics indicated that the information derived from benchmark assessments were of limited use in providing valid and reliable results. Further, qualitative survey results (See Appendix D.) as well as teacher and principal interviews revealed that the time devoted to testing and accountability initiatives limits quality instructional time, impacts the depth of curriculum coverage, interferes with the development of students' critical and divergent thinking skills, consumes a significant amount of district monies, and results in high school graduates being less prepared to function independently as young adults in society.

ARE BENCHMARK ASSESSMENTS USEFUL MEASURES?

While the project focused on identifying areas of efficiency and effectiveness at the campus and district level in near real-time, initial quantitative analyses examined the practical utility of current benchmarks, as well as the relation between benchmark assessments and TAKS scores. Overall, the results were mixed.

- Reading/ELA benchmark assessments were ***moderately strong*** indicators of TAKS reading performance among all middle schools
- Overall, Algegra1 benchmark assessments were indicative of TAKS math performance. However, among the schools receiving lower efficiency and effectiveness scores, the results were mixed.
- The correlation between eighth grade math benchmark assessments and TAKS math performance were mixed among all middle schools.

The results of the quantitative analyses revealed that the reading/ELA benchmark assessments were indicative of student performance on TAKS reading . Correlations between reading and English language arts benchmark assessments and TAKS reading performance among regular education students ranged from r = .236 ($p < .01$) to r = .377 ($p < .01$), among all middle schools. The results of the regression analysis indicated benchmark assessments were statistically significant predictors of TAKS reading performance, net the effects of student demographic variables. The assessments explained approximately 23% of the variance collectively in the outcome variable, which, according to Coehn (1988) is a moderate effect. Benchmark 2, which was based on reading skills appeared to be the best predictor of TAKS reading performance as evidenced by the standardized beta coefficient (β = .31), while the initial benchmark focusing on English Language Arts content was the least important predictor. Similar results were found among each of the five middle schools participating in the qualitative analysis.

Regarding mathematics performance, statistically significant correlations among all middles schools ranged from r = .441 to r = .692. ($p < .01$). Separate regression

analyses were conducted for the Algebra 1 and 8th grade math benchmark assessments. Regression results related to Algebra 1 assessments indicated that each assessment was a statistically significant predictor of TAKS math performance ($p < .01$). Each one-point increase on the initial Algebra 1benchmark assessment was associated with an increase of 1.09 scale score points on TAKS math. Similarly, each one-point increase on the Algebra 1 benchmark 2 and 3 assessments was associated with an increase of 2.21 and 1.44 scale score points on TAKS math respectively, net the effects of the student demographic variables. While the overall results indicated that the Algebra 1 benchmarks were statistically significant predictors of TAKS math performance, results were mixed among the five middle schools selected for the focus group analysis. At H.E. Charles Middle School, none of the Algebra 1 benchmark assessments were statistically significant predictors of TAKS math, while only the second benchmark assessment was a statistically significant predictor in Bassett and Canyon Hills Middle Schools. Note while the results were somewhat random, not all schools participated in each benchmark administration.

Concerning the relationship between 8th grade math benchmark assessments and TAKS math performance, results revealed benchmark 2 was not a statistically significant predictor of TAKS math performance ($p = .26$), while benchmark 1 and 3, were statistically significant ($p < .01$). The standardized beta coefficients indicated that benchmark 3 was the best predictor of TAKS math performance among each of the variables examined (β = .32), while the standardized beta coefficient associated with benchmark 1 was .24. Each one-point increase on benchmark 3 was associated with an increase of 1.57 scale score points, while each one-point increase on benchmark 1 was associated with an increase of 1.20 scale score points on TAKS math. While the overall results indicated that the first and third benchmark assessments were statistically significant predictors of TAKS math, the results among the five middle schools selected for the focus group analysis were mixed. Among the five schools visited, benchmark 1 was not a statistically significant predictor of TAKS math in Bassett Middle School. At Canyon Hills, both benchmarks 1 and 2 were statistically significant predictors of TAKS math performance, while benchmark 2 and 3 were statistically significant predictors of performance on the outcome variable at Lincoln and H. E. Charles Middle Schools. A plausible explanation for the mixed results found in individual schools could be attributed to the utility and quality of the benchmark assessments. The majority of teachers interviewed believed that the assessments were not aligned with the pacing guide and therefore did not value nor fully utilize the results of the assessments to alter classroom instruction. In addition, teachers reported concerns about the quality of the assessments. To better understand the underlying properties of the assessments, the authors conducted an item response analysis among the assessments. The results of the analysis are reported below.

The quality of the benchmark assessments among the middle schools examined is marginal at best. In the following example, a 1 factor solution was determined to meet the assumptions of the item response model calculated. The overall reliability of the factor was .60, which is below the commonly accepted threshold of .70 for common assessments and .90 for high-stakes assessments. In addition to the low reliability, the assessments appear to be limited in difficulty while providing no information about the knowledge and understanding of the subject matter among students with upper-level ability (i.e., the questions are too easy with little discrimination between those that know and do not know the subject matter). The information curve depicted in the following figure is representative of the

benchmark assessments in mathematics in the middle schools during the timeframe of the study. Note the results depicted in Figure 5 were from an 8th grade math benchmark assessment. To further explain, the information curve of a properly constructed assessment should be centered above zero. In the following figure, the information curve is centered at approximately -1.98.

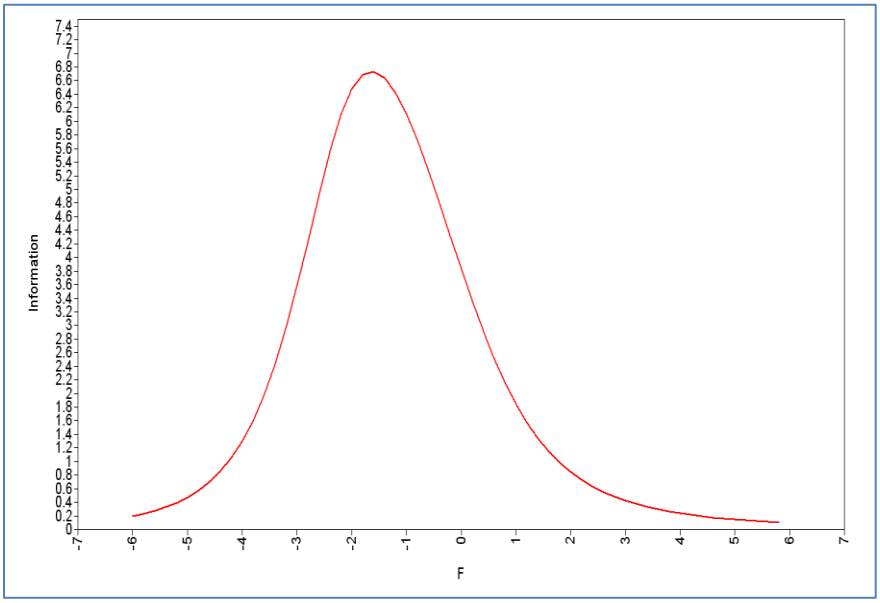

FIGURE 5. TEST INFORMATION CURVE FOR A REPRESENTATIVE 8TH GRADE MATH BENCHMARK ASSESSMENT

HOW MUCH TIME IS LOST?

Teachers consistently expressed significant concern about the amount of instructional time that was lost and subsequently devoted to current accountability initiatives in terms of preparing for, administering, and reviewing the following: district benchmarks, common assessments, mock TAKS, and TAKS. Math, science, social studies and ELA teachers reported losing the most amount of time due to district accountability initiatives. Each of these content area teachers reported losing 3 periods simply to the administration of benchmarks, up to 4 days to the administration of TAKS, and a minimum of 36 days to the administration of common assessments in the 2010-2011 academic year. Assuming a 180 calendar year and assuming all classes meet daily, this equated to students physically taking exams for 43 days, or roughly 25% of their school year, in each content area. Additional instructional time was lost for

> *"If we could get rid of benchmarks and common assessments and were allowed to follow teaching guides we would have better results at the end of the year because we would have more instructional time, and not as much testing time. And it would be good for EVERY student. I don't mind the tests because kids will have to test when they get into their careers…but we are testing too much."*]

individual students in schools where remediation was done by pulling students out of non-TAKS tested subjects such as science and social studies.

Although 91% of online survey respondents indicated they used assessment to improve instructional practices, teachers interviewed stated that too much time was wasted on data collection and analysis. In fact, teachers reported activities such as exam preparation/tutoring both in and outside school hours; the examination of item analyses during teacher planning time; and in-class reviews of exam results consumed additional and even more extensive amounts of their instructional and planning time. Instructional coaches noted that their availability for teacher improvement, including modeling instruction, was severely limited because of the time dedicated to these assessment activities.

Keeping Up The Pace

Teachers indicated they were expected to cover district curriculum at a pre-determined pace. As a result, "Gifted and Talented" (GT) teachers reported feeling restricted at times by district curriculum, whereas teachers working with inclusion and LEP populations reported feeling torn between ensuring students had developed appropriate content mastery and moving forward to keep up with district curriculum. Many teachers felt their campus principals personally valued their expertise and afforded them great academic freedom so long as students were passing TAKS and campuses were meeting AYP requirements. In fact, 81% of online survey respondents indicated their campus administrators included them when developing campus policies, objectives, missions and goals. However, these same teachers perceived feeling pressure from the district central office to progress through curriculum on a pre-determined timeline before students had adequately mastered content.

"We're missing two and half hours of instructional time each week this year and the common assessments take up another hour each week. Our principals listened to our concerns about this and offered suggestions, even though they couldn't really change the common assessment rules. The district requires a lot of things of us that the campus administrators cannot control. We have the same curriculum and the same expectations as we had before, but with less instructional time each week."

"Some teachers and students feel rushed because we do those weekly assessments and if the students are behind we know we have to re-teach, and then we fall behind when we do that. The students and teachers sometimes feel overwhelmed at times."

Teachers reported accountability and scheduling challenges led directly to the loss of instructional time. To illustrate, some teachers reported that scheduling changes at their campuses had resulted in the daily loss of 30 minutes of instructional time during the 2010-2011 academic year. This totaled a

loss of 90 instructional hours for the year, in addition to the loss of instructional time to testing. Teachers reported the district continued to expect them to cover the same amount of content during the school year as they were expected to cover before these scheduling changes were set in motion.

Strength of Character and Critical Thinking Skills: Casualties of War

Many teachers shared "back in the day" stories when talking about how current students were different from those who walked through school doors ten or even twenty years ago. Teachers associated with this study described a successful student as one who studied, came prepared for class, did his/her homework, improved over time, and displayed dignity and confidence. Less often did teachers report successful students as those who passed required TAKS examinations, though teachers and administrators from each campus indicated that being able to pass TAKS was an important milestone required to progress on to the next grade level.

> *"The district stresses the TAKS as a sign of success, and students come here and that's what they see, so that's what they learn. Years ago we were teachers, but now we're test taking strategy teachers. We don't have time to teach them how to think for themselves. If it's not on the test, we don't even cover it. Right now the students are not well rounded. By the time they get out high school, their heads are just shaking. I really blame these standardized tests on our high school graduates not being able to function in society when they get out of school."*

However, teachers reported there had been a paradigm shift in terms of the way students approached and viewed school over the last decade. Teachers indicated the loss of quality instructional time, enhanced by the anxiety associated with test preparation and performance, had contributed to students becoming skilled test takers, but not critical or divergent thinkers. Teachers expressed concern that high stakes testing initiatives had changed the academic culture of their schools in such a way that students were not adequately prepared to function as responsible citizens beyond middle school and into early adulthood.

The online survey also examined changes in academic culture in light of high stakes testing mandates, but results were mixed. Only 75% of respondents indicated that today's students would be likely to engage in more challenging work if given the opportunity, though 94% of respondents indicated they expected their students to become independent learners.

A Penny Spent is A Penny Lost

For more than two decades the education reform movement has focused on the academic output of schools. This attention has resulted in many new educational initiatives, such as those with greater emphasis on math and science, those which prepare more students to be ready for college, and those which increase the quality of teacher and principal effectiveness. While this academic focus is not diminishing, a new consideration is dominating the reform movement spotlight: financial accountability. This new focus is timely as Texas faces one of the largest budget crises in the state's history where spending on public education -- the largest category in the state budget -- has nearly doubled to $55 billion since the 1998-99 school year, with spending per student increasing 63% (State Comptroller, 2010).

- EPISD incurs roughly $32,000 each day in operating costs in order to keep campus and administrative doors open
- EPISD spends roughly $300,000 annually to print district benchmarks exams. The cost of instructional time lost and printing expenses totals $1.4 million annually.

While recent studies have compared schools and districts across the entire state (State Comptroller, 2011), using other districts as a comparison may not be the best method of designing school improvement as each district has unique characteristics that may preclude a fair comparison. Some of these unique district characteristics may include the economies surrounding districts, variability in student demographics, and differences in both culture and values. Because of this the EDC project examined the financial efficiency of each EPISD campus in producing student outcomes in order to "get their own house in order." Comparing campuses within EPISD was advantageous on a number of levels: comparing like campuses allowed for a comparison group that was readily accessible, the goals and objectives of each campus were similar, dollars expended were on the same scale, and student demographics and community values and culture were more similar within EPISD than between districts. **This grass roots approach yielded an estimated potential savings of more than $11 million among middles schools alone, in the areas of instruction, school leadership, instructional resources and media, guidance and counseling, and campus maintenance and operations**. The process for calculating these savings and results of analyses are reported below.

The longitudinal comparison of efficiency scores across a three-year period among all middle schools is displayed in Table 1. The results revealed that the efficiency scores varied greatly among H. E. Charles and Bassett Middle Schools. These two middle schools showed significant gains from 2008 to 2010. In contrast, Ammedariz, Guillen, and Wiggs Middle Schools were consistently the most inefficient campuses during the three-year timeframe. Finally, Brown Middle School was the only campus that operated at an efficient level in each of the years examined (Mean Efficiency = 100, SD = 0).

TABLE 1. EFFICIENCY SCORES OVER A THREE-YEAR TIMEFRAME

Campus	Efficiency Score 2008	Efficiency Score 2009	Efficiency Score 2010	Average Efficiency Score (3 Years)	Std. Dev. Efficiency Score (3 Years)
Armendariz MS	65.69	64.72	61.85	64.09	2.00
Bassett MS	46.52	58.50	68.89	57.97	11.19
Brown MS	100.00	100.00	100.00	100.00	0.00
Canyon Hills MS	74.54	83.00	80.67	79.40	4.37
Charles MS	68.63	100.00	80.42	83.02	15.85
Guillen MS	60.99	56.88	61.43	59.77	2.51
Henderson MS	59.39	68.96	60.19	62.85	5.31
Hornedo MS	100.00	100.00	100.00	100.00	0.00
Lincoln MS	77.33	100.00	100.00	92.44	13.09
Magoffin MS	58.63	65.94	72.52	65.70	6.95
Morehead MS	69.99	81.07	77.88	76.31	5.70
Richardson MS	79.89	89.95	86.84	85.56	5.15
Ross MS	73.58	81.54	71.62	75.58	5.25
Terrace-Hills MS	57.49	81.66	64.10	67.75	12.49
Wiggs MS	74.88	79.47	76.30	76.88	2.35
Grand Mean	71.17	80.78	78.00	76.65	4.95

POTENTIAL SAVINGS

Potential savings were examined across five functions (instruction, school leadership, instructional resources and media, guidance and counseling, and campus maintenance) in three key areas that included General Revenue (fund 199), State Compensatory Education (fund 185), and Title 1 – part A (fund 211). This phase of the analysis focused on the most recent budget data available (fiscal year 2010). Regarding General Revenue, the results displayed in Figure 6 (controlling for campus demographics) indicated that to achieve maximum efficiency, a reduction of 16% in instructional expenditures, approximately 22% in instructional resources and media, 19% in school leadership, 18% in guidance and counseling, and 25% in campus maintenance and operations was necessary.

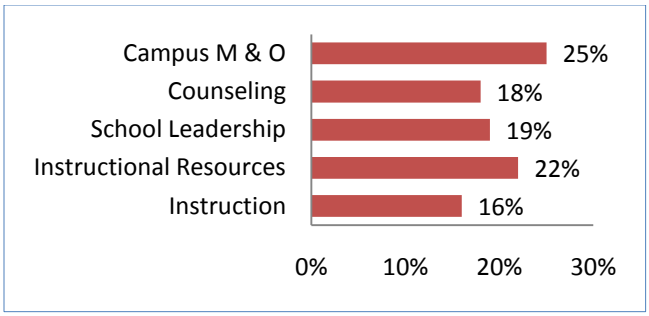

FIGURE 6. PERCENT REDUCTION NECESSARY FOR ALL MIDDLE SCHOOLS TO ACHIEVE 100% EFFICIENCY. NOTE EXPENDITURES BASED ON FUND 199 (GENERAL REVENUE) ONLY.

The majority of potential savings were recommended in the area of instruction where more than 70% of total dollars expended were on teacher salaries. Note EPISD spends approximately $50,000.00 annually on individual teacher salaries. To achieve efficiency in this area, it was recommended that EPISD examine student and teacher schedules among all middle schools to determine if the number of teachers could be reduced. Schedules varied from a 7 period day to a 6 by 6 block schedule across campuses. While a 7 period day requires fewer teachers, a 6 x 6 block schedule allows for more time on task and in-depth instruction. While achieving a balance between effectiveness and efficiency is challenging, it is a challenge that must be addressed to ensure an effective instructional program in the most efficient manner.

Regarding State Compensatory Education, a reduction of 15.7% in instructional expenditures and 21.7% in staff development was necessary to achieve maximum efficiency. While the majority of potential savings was in the area of instruction, the area of staff development should not discounted. Based on in-depth analysis, past staff development had little impact on teacher effectiveness.

Finally, regarding Title 1 funds, a reduction of 32.7% in instruction, 19.8% in staff development, and 13.1% in school leadership was necessary to achieve efficiency and maximum student performance. While it may not be possible to reduce spending in this area due to the nature of the fund, the results suggested the percent reduction should be reallocated to effective instructional programs to ensure that disadvantaged students achieve maximum performance on intended outcomes. As evidenced by in-depth analyses and interviews with key administrators, many schools had a substantial sum of Title 1 dollars that had to be expended each year. At the end of each school year, if the money is not expended, schools lose those monies. Therefore, given that the money must be spent within predetermined time frames, many campuses reported purchasing items that may or may not be directly linked to improving student performance. Spending money in a potentially frivolous manner can have a significant negative impact on efficiency and instructional effectiveness among students who need the most assistance.

Looking Ahead

While results of the current study were based on static data at the fund and function, and object level, combining the DEA and SFA results with regression procedures and examining line-item expenditures could allow campus and district personnel to pin-point specific expenditures for potential savings. While the analytics are complex, results can be generated and presented in an easy-to-understand format. From a preventive stance, campus principals and district personnel could monitor a dashboard in near real-time that shows when a campus or district goes from efficient to inefficient (Figure 7). Having access to data in a useable format in near real-time has the potential of ensuring that districts are getting the most out of each dollar expended.

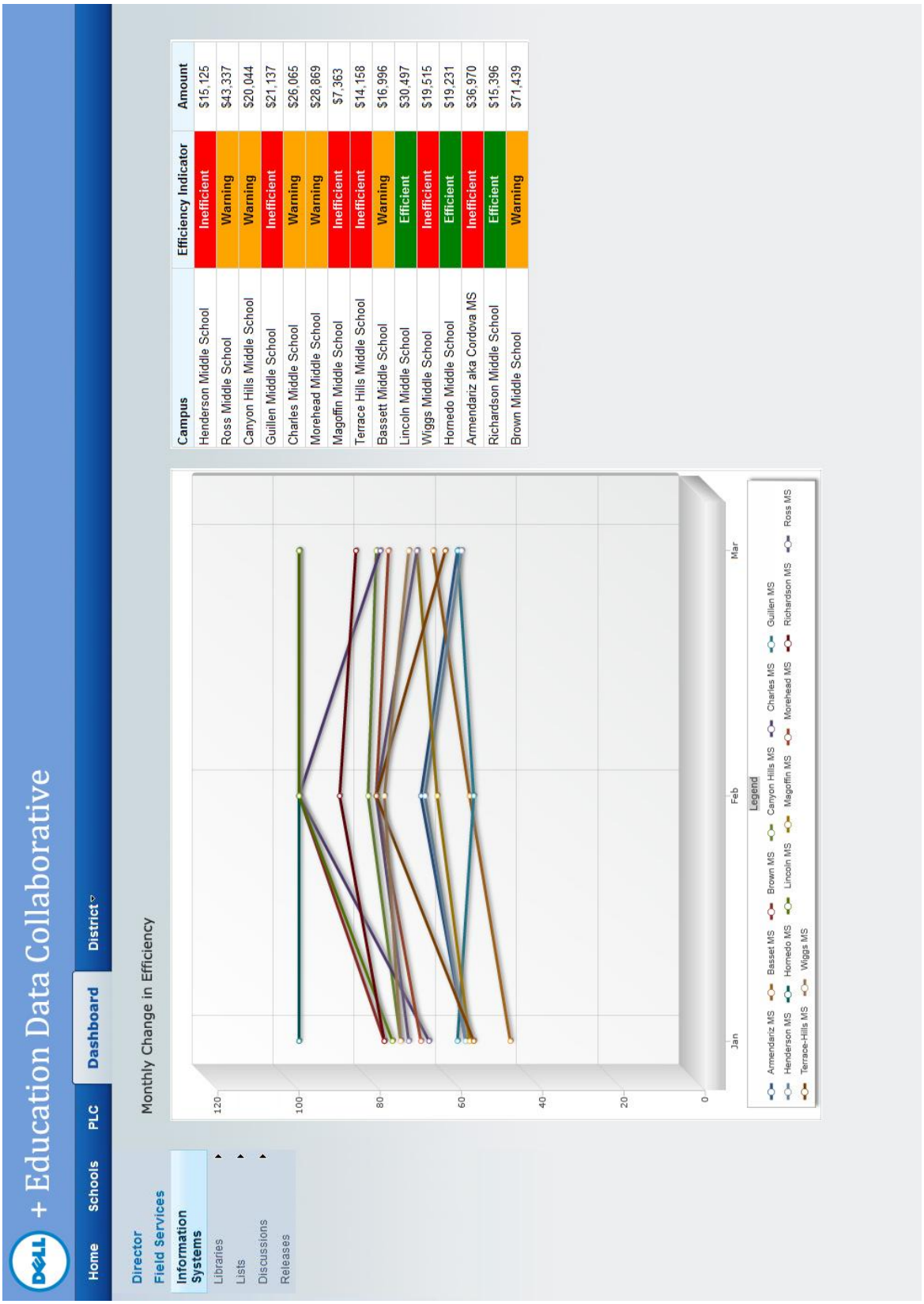

FIGURE 7. MONTHLY CHANGE IN EFFICIENCY

Teacher Effectiveness: Can Everyone Be Great?

In this study, teachers and administrators indicated teachers were formally evaluated on the basis of classroom Walk-Throughs by school administration, the use of the Professional Development and Appraisal System (PDAS). Nearly all teachers agreed they were also informally evaluated by administration on the basis of their students' TAKS performance. Teachers stated that Walk-Throughs took place on a regular basis, as all campus principals and assistant principals conducted ten weekly walk-through evaluations of teachers. These evaluations were reported as less formal in nature, unannounced, and generally spanned 10-15 minutes in duration.

In contrast, teachers indicated the formal PDAS evaluations were announced and scheduled ***with them*** and generally spanned 45-60 minutes in length. Interview data revealed beginning teachers were evaluated once per academic year during years 1-3, but could request a waiver from PDAS evaluation during year 3. Further, it was reported that teachers who are new to campus but come with more than three years experience only get evaluated via the PDAS instrument once at the end of the third year in their new school. Teachers and principals indicated these infrequent evaluations as the basis for retention decisions. Elements of PDAS evaluations included:

- Active, Successful Student Participation in the Learning Process
- Learner-Centered Instruction
- Evaluation and Feedback on Student Progress
- Management of Student Discipline, Instructional Strategies, Time, and Materials
- Discipline procedures

Teachers and principals indicated TAKS assessments are administered during the spring of each academic year on dates set forth by the state, usually in March and April. Eighty-five percent of online survey respondents reported their success as measured by growth in student achievement. Similarly, 91% of online survey respondents indicated assessment data was used to evaluate teacher effectiveness. Table 2 below outlines current state mandates regarding required TAKS by grade level.

Table 2. Middle School State TAKS Mandates by Grade Level

6th Grade	7th Grade	8th Grade
Reading (English or Spanish)	Reading	Reading*
-----	Writing	-----
Math (English or Spanish)	Math	Math*
-----	-----	Science
-----	-----	Social Studies

* Beginning in 2008, students are expected to pass this exam in order to progress to the 9th grade.

Beyond PDAS and Walk-Through evaluations, teacher effectiveness was examined further in this study and on the basis of TAKS and benchmark scores (see "Quantitative Procedure: Effectiveness"). Findings from resulting teacher effectiveness analyses indicated there was little variation in teacher effectiveness scores among campuses that performed above the median on TAKS math and reading, and that math teachers maintained higher teacher effectiveness scores than reading teachers. Among campuses scoring below the median in these same subjects, teacher effectiveness varied significantly, especially in reading, while teacher effectiveness scores among math teachers were rather stable. Teacher experience appeared to be strongly associated with teacher effectiveness in the lower performing schools. In fact, reading teachers with 10 or more years experience maintained lower teacher effectiveness scores than teachers with 3-7 years experience.

- There was little variation in teacher effectiveness scores among campuses that performed above the median on TAKS math and reading
- Among campuses scoring below the median in TAKS math and reading, teacher effectiveness varied significantly, especially in reading, while teacher effectiveness scores among math teachers were rather stable.
- Teachers with 10 or more years' experience maintained lower teacher effectiveness scores compared to teachers with 3-7 years' experience

Teacher Perceptions Regarding Walk-Throughs, PDAS and TAKS

Teachers were in near uniform agreement regarding perceptions of the strengths and weaknesses of PDAS and Walk-Through evaluations. Teachers reported Walk-Throughs as more authentic measures of teacher effectiveness than PDAS evaluations. Because Walk-Throughs were reported as unannounced evaluations that only take place for 10-15 minutes during any given day, teachers indicated these evaluations provided administration with frequent, natural "snap-shots" of their teaching.

"Administrators come in one or two times to evaluate us and they expect to see a 'dog and pony show.' We do this to get the marks we want to get, and it's all so superficial. They know we are good teachers, but they have to do these evaluations. I think a Walk Through is better than having to sit through a whole lesson. It's more authentic and it's less stressful and it's not as long. Our principals already know, through their years of experience, how well the teachers work with the students."

PDAS evaluations, on the other hand, were perceived as inaccurate measures of teacher effectiveness because they normally occurred only once per academic year. Described as "Dog and Pony shows," PDAS evaluation dates were reported as announced, and teachers indicated they took extra time to ensure they prepared lesson plans which incorporated all PDAS elements. Further, teachers readily indicated those lessons were not entirely reflective of how they would normally teach. In fact, no teacher expressed confidence in PDAS as an accurate or fair evaluation tool for their teaching.

Across study schools, teachers stated student performance on TAKS assessments was closely monitored by campus principals and teachers. Teachers indicated TAKS scores as an important aspect administration considered when evaluating teachers, though they uniformly did not describe TAKS as a valid measure of student achievement or teacher effectiveness. However, teachers indicated administration carefully considered more subjective evaluations of teacher effectiveness, such as the Walk-Through evaluations. Teachers consistently reported administration as also concerned with elements such as whether or not students were actively engaged in learning, the extent to which teachers were directly involved with the students during instructional time, and the clarity of lesson objectives.

Principal Perceptions Regarding Walk-Throughs, PDAS and TAKS

Principal perceptions mirrored those of teachers as they considered Walk-Throughs more accurate measures of teacher effectiveness in the classroom, but principals did accept PDAS as a long-time requirement for teacher evaluation in Texas. Principals stated that although Walk-Throughs took place throughout the school year, PDAS evaluations generally only took place in the spring semester. The fact that principals elected to wait until spring to begin PDAS evaluations further illustrated the lack of buy-in among administrators that PDAS was a reliable, authentic tool for evaluating teachers. TAKS outcomes were another piece principals stated they informally considered, but they indicated PDAS and Walk-Through evaluations were most often utilized as venues for providing teachers with praise and/or constructive feedback which is intended to improve teaching effectiveness.

When asked about what makes a successful teacher, one principal stated, "Our teacher of the year is one of the most caring people I've ever met. He truly cares about the student and has the ability to drive them because he has such an ability to put fun things together that are challenging. He has a 100% pass rate on the TAKS. A successful teacher really shows he/she cares about his/her students."

When asked about how to work with teachers who have many years of experience, but appear to be ineffective, a principal responded, "You just have to build on the culture that you inherited…some days it's very tiring, but all you have to do is walk outside and see the kids, and then you know why. If I'm expecting my teachers or staff to do something, I have to be willing to do the same thing."

When asked if TAKS, PDAS, and Walk-Through evaluations were utilized as evidence to support the removal or reassignment of ineffective teachers, principals indicated that EPISD did not possess a culture which actively removed or reassigned ineffective teachers, especially if they had worked in the district for many years. Although not all principals indicated buy-in to this culture, there was generally a low willingness to use TAKS, PDAS and Walk-Throughs to identify and possibly to remove ineffective teachers.

LOOKING AHEAD

While school and district leaders strive to have a quality teacher in every classroom, it is common practice to wait until the end of the school year, and after the release of TAKS or the State of Texas Assessments of Academic Readiness (STAAR) results, to determine the instructional impact teachers had on their students. The results of the current pilot study revealed this, too, as EPISD district policy. However, in the EDC project, it was determined that teacher effectiveness could be monitored in near real-time throughout the current school year. Using a specially derived longitudinal model and well-designed, vertically linked assessments, principals could monitor student growth or progress and teacher effectiveness simultaneously in 6 or 9 week intervals. From a preventative stance, school and district leaders could monitor each teacher's effectiveness score, which is regularly updated throughout the school year, from a dashboard that provides in-depth performance data about each teacher (Figure 8). Based on near real-time monitoring of teacher performance results, principals could immediately determine areas of need for individual teachers and provide meaningful professional development to address identified areas when a problem is identified ***during the school year***.

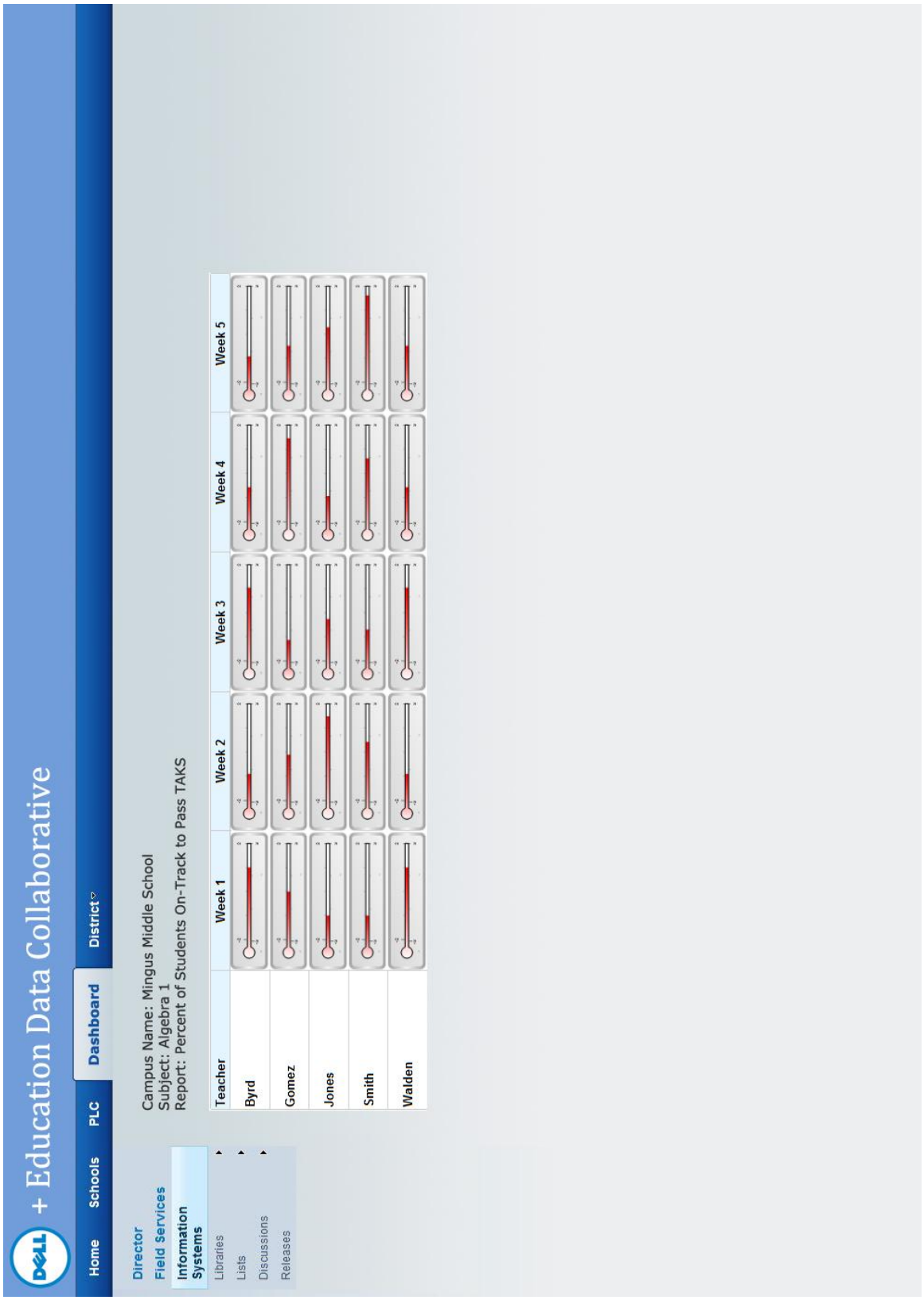

FIGURE 8. CAMPUS ON-TRACK DASHBOARD BY SUBJECT

Cultivating an Effective School Climate

Principals reported utilizing proactive measures for ensuring effective teachers instead of relying on the reactive measures of removing or reassigning ineffective teachers. Some of these proactive measures included formally and informally assigning mentors to new teachers. Although only 55% of online survey respondents indicated it would have been helpful to have had their education professors mentor them as new teachers, almost all interviewed teachers shared that campus-provided mentors made a significant, positive impact on their effectiveness as new teachers. Other proactive measures campuses reported to employ included actively seeking to hire teachers who naturally possess a genuine concern for children, actively seeking input from teachers regarding scheduling, maintaining high expectations of teachers, affording teachers as much academic freedom as is reasonable, and developing a sense of community within their schools where teachers feel valued and empowered.

When asked about how a principal built a school culture which cares about children before caring about student outcomes, a principal responded, "You have to look for applicants who already have that in them. I realize that anyone that I hire has to possess a natural concern for children. Knowing how to deliver content will come later, as you can mentor them on delivering content. We don't have many teachers who leave this campus. It's been that culture always."

PROFESSIONAL DEVELOPMENT: LOTS OF ARROWS - NO CLEAR TARGET

- Recent staff development had no significant impact on teacher effectiveness
- Professional Learning Communities (PLCs) were an effective structure for building collaboration and teacher sharing
- Effective PLCs showed the critical importance of quality leadership

PROFESSIONAL DEVELOPMENT

EPISD's commitment to professional development has been considerable. Professional development (PD) was reported as offered on an ongoing basis through building-based programs, district level staff provided training, outside consultants, and a limited number of attendees at state and national level conferences. In 2009-2010 over $28 million[4] was spent to improve classroom instruction, advance leadership initiatives, and expand the skill and capacity of staff to improve student performance district wide. On the conservative side, at least $6 million could be viewed as being targeted towards middle school improvement.

To analyze the effect of professional development, teacher effectiveness scores were generated for all teachers in both mathematics and reading among all EPISD middle school teachers. These effectiveness scores were subsequently compared with the number of days teachers participated in professional development activities.

Results revealed a statistically significant correlation between the number of staff development days in which math teachers participated and their teacher effectiveness score ($r = .236$, $p < .01$). Note that while the results are statistically significant, the effect size associated with the increased number of staff development days and improved teacher effectiveness scores is only .06, indicating that the increased number of staff development days explain approximately six percent of the variance in teacher effectiveness, while 94% is left unexplained. Similar results were found among reading teachers where the correlation was ($r = .241$, $p < .01$), with an effect size of .048, indicating that the number of professional development days attended explained less than five percent of the variance in teacher effectiveness (Table 3).

[4] Function 13, AEIS Report

TABLE 3. PEARSON PRODUCT-BETWEEN TEACHER EFFECTIVENESS SCORES, NUMBER OF STAFF DEVELOPMENT DAYS ATTENDED, AND OTHER SELECTED TEACHER CLASSROOM ABSENT VARIABLES

Teacher Effectiveness (1)	1.00						
Staff Development (2)	.236**	1.00					
Personal Illness (3)	-.153	-.128**	1.00				
Family Emergency Leave (4)	-.156	-.035**	.056**	1.00			
Personal Business (5)	.079	-.130**	.193**	.093**	1.00		
Absent on School Trips (6)	.097	.036**	-.062**	-.022*	-.044**	1.00	
Military Duty Leave (7)	-.008	-.008	-.002	-.003	.001	.003	1.00

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Thus, although there was a statistically significant relationship between professional development and teacher effectiveness, professional development activities, when considered in isolation, had little direct impact on teacher effectiveness thus calling into question the quality of the professional development offered and the efficient use of funds expended on professional development opportunities. The resulting unexplained variance indicated there were other mediating variables, in addition to professional development, which were correlated with and/or influenced teacher effectiveness. This finding supported qualitative interview results, where teachers indicated their effectiveness was enhanced not only by the number of days they participated in professional development experiences, but also by the quality of professional development offered, the types of PLC planning opportunities they had, their relationships with students, their administrative support, and the extent to which academic and instructional resources were available.

THE EXPERIENCE OF PROFESSIONAL DEVELOPMENT IN EPISD

Despite the lack of evidence to support a direct link between professional development and student performance, can the experience of engaging in professional development be a positive factor for teacher morale and persistence and, thus, have an indirect effect on student performance?

Eighty five percent of the teachers responded positively when surveyed about professional development improving instruction in the classroom. Yet, in schools that showed significant improvement over the last five years, no teacher identified the professional development they received as a contributing factor to what made their school successful. Overall, teachers perceived some of the consultant delivered PD to be useful, but most of the district professional development was reported by teachers as ineffective. Three specific professional development strands consistently emerged as being effective such as Kagan Strategies, Differentiated Instruction (Dr. Bender), and Laying the Foundation. More may have emerged if additional probing was focused on less recent history. Teachers who

spoke highly of the professional development noted the intrinsic motivation of the teacher to participate. This factor helped explain the inconsistent responses in the focus groups.

Teachers and principals reported the district has moved over the last few years towards a "train the trainer" model. Instructional coaches and other leaders are trained and then expected to bring the knowledge to their respective buildings.

Quote from teacher: "*I'm a coach, so I get PD monthly… but I don't have the time to bring it back to the school. If something is free or if a sub is provided by the district, then admin is highly motivated to participate. The district can send facilitators to the PLCs, but when that happens it often becomes a gripe session and they aren't productive.*"

Teachers and school leaders noted challenges with this model to include: the lack of fidelity to the original PD sessions; the concentration of time, such as a two day training expected to be delivered in 1-2 hours; and the lack of expertise of coaches to answer questions or access resources they did not experience in their session. Others experienced a strong disconnect between the goals of the PD and their resources for teacher and learning. Compounding these concerns were the perceived financial limitations of accessing professional development.

Schools with Title 1 funds indicated they were able to fund substitutes relatively easily when compared to their non Title 1 peers. Further, specialized programs that were subsidized with separate funding streams indicated securing funds needed to support PD activities was more easily accomplished. GT and Special Education teachers noted that they were pulled out of their classes frequently compared to their peers. Relatively new and veteran teachers noted that the best professional development for inducting new teachers had been helpful, but the drop off in quality PD needed to sustain those practices diluted the effect. When the district provided substitute funding there was greater support from building principals for the training.

Quote from teacher: "*We have too much PD and it's not of high quality. We'd rather be in the classroom. 90% is useless and 10% is good stuff maybe, 80/20… you have to hunt for the good stuff. Sometimes we get assigned staff development, like in technology. Well we don't have the technology at our school, so why are we going to technology PD?*"

Finally, outside factors were reported to contribute to the effectiveness of PD. Many teachers noted a lower quality of instruction was delivered when they were away from the classroom despite their best efforts to leave meaningful assignments. The lack of training and supervision of substitutes was reported to contribute to an academic culture that suggests it is more important for a teacher to be in his/her classroom than it is for that teacher to get help on improving. Scheduling alternatives such as after school, Saturdays, and during the summer were encouraged by some, but not by a majority of teachers as a desirable option.

Linking Professional Development to Supervision

As described before, principals and teachers indicated EPISD administration currently does not have the ability to systemize their teacher development model. Neither PDAS nor walkthrough data were reported to align to specific staff development programs that could improve teacher performance in response to administrative intervention. Clear, actionable patterns did not emerge through the walkthrough data because administrators do not use it as a supervisory tool linked to evaluation. Administrators commented that if they saw something in a classroom that concerned them during a walkthrough they typically would speak with the teacher about it. In rare cases did these conversations become part of any teacher's performance record. As such, staff development was not based on areas of weakness identified through the supervision process, but rather as informal recommendations.

> *"Without a clear supervision and evaluation model to provide actionable data on specific areas of need, internal professional development offerings are ultimately tied to a broad goal of improving performance on assessments with insufficient differentiation to teacher need."*

> *"In the beginning I was resistant of the PLCs, but it's helped us to get to know each other and our experiences, and we're able to keep up with individual students. We're able to have an opportunity to plan together and that's been extremely helpful".*

Further, teachers did not connect professional development with the observation process. When asked about how professional development offerings are targeted, most teachers noted the connection to broader school or district goals. Rather than "owning" the responsibility for PD because it is part of an individualized instructional improvement plan, teachers experienced a disconnect between what they want to learned about, what they were expected to attend, and what should be the resulting impact on their effectiveness. Without a clear supervision and evaluation model to provide actionable data on specific areas of need, teachers and principals reported that internal professional development offerings were ultimately tied to a broad goal of improving performance on assessments with insufficient differentiation to teacher need.

School Based PD: Professional Learning Communities

School level, professional learning communities (PLCs) were reported to have substantially improved the culture of collaboration across all of the schools; however, the dramatic inconsistency of time, leadership, and topic focus across schools suggested that PLCs need to be better understood and utilized more often as an effective and efficient model to improve instruction.

In the higher performing schools, teachers viewed PLCs as a helping their teaching and learning. Teachers in lower performing schools did not report consistent positive experiences with PLCs. Further, teachers in the higher performing schools noted a clearer focus in PLC activities. These teachers described percentages of time spent on various activities (data analysis, lesson planning, and modeling of lessons) consistently across focus groups. In schools where the attitudes and perceptions towards PLCs was less positive, there was significantly more disagreement in articulating how time in the PLCs was spent. Interestingly, also in higher performing schools, teachers described their PLC time as "professional development time" more frequently as well.

> *"The best is during PLC…it's probably the most effective and efficient way for you to learn what's important for your students, your school and your content area. The district training is general for all, not for just your classroom."*

The implementation of PLCs varied widely. In some schools, PLC time was reportedly built into the schedule every day, while others indicated PLCs met two or three times a week. Some campuses stated PLCs met only one day a week. In terms of activities associated with PLC time, some campuses noted spending nearly 100% of their time analyzing data while other campuses indicated they spent very little time analyzing data. Some met by subject area across the three grades, while others met in large groups across content areas. No one pattern emerged across the study schools. The key drivers in the success of these multiple configurations appeared to be the attitudes of teachers towards their PLC as a professional development experience and the PLC leader.

EQUITY – STUDENT TEACHER RATIOS, STUDENT LOAD, AND ACCESS TO THE BEST TEACHERS

- Lower student/teacher ratios may not lead to improved TAKS scores
- Student/teacher ratios in core subjects much were higher than student/teacher ratios for specialized classes such as GT
- The best teachers may not be with the students who need the best instruction

In each school, intervention programs for mandated remediation were reported as firmly established. Principals and teachers indicated that, each spring, students in need of additional instructional are identified using TAKS performance data, grades, and local factors. This group of students then receives additional instructional time in English/Language Arts and/or math during the school day in place of elective opportunities. Also, as mandated by state and federal law, sub population groups of Limited English Proficiency (LEP) students and Special Education students were reported to receive additional instructional services beyond their mainstream classes to provide scaffolds to the curriculum. Teachers and principals at all campuses indicated these two sub groups were carefully monitored for progress towards goals.

It was reported that many special education students mainstream in core classes described as "inclusion" classes. Teachers stated these inclusion classes were composed of both targeted students who received necessary services in the classroom and regular education students. Teachers indicated these classes were typically taught by two adults: the content area teacher and a special education teacher. Further, a goal of this teaching team was to deliver instruction to all students in a manner where there was no discernable difference between regular education and special education students.

Similarly, GT classes and programs were reported to provide an alternative and more challenging curriculum for students who demonstrate higher academic skill relative to their peer group. Teachers indicated there was no additional staffing provided for these classes, although at many campuses modifications were made to accommodate combined sections of ELA and social studies into a single humanities class.

Traditionally, the effectiveness and efficiency model analysis tended to examine the performance levels of students in specific programs or for sub-populations because they were part of the accountability system. However, a full analysis should also attempt to understand the effect these divisions have had on the full school population in order to identify to what degree EPISD, as a whole, has helped its students reach their academic achievement potentials.

SKIMMING THE TOP

Typically a small percentage (up to 5%) of students is classified as GT to meet the state mandate[5]. Schools have the option of adding students to these classes based on local criteria. EPISD reported consistently accommodating additional students in these classes based on the criteria of grades, performance on

[5] Section 29.123 of the Texas Education Code

standardized tests, and teacher recommendation. The relative number of additional students varied by campus as did the percentage of overall students. At a majority of the study schools, this percentage was reported to be about 15%.

Ninety-one percent of the study school teachers indicated they have high expectations for students. Research has indicated that this belief is a critical piece to overall school performance. Additionally, 75% of the surveyed teachers believed that "if students are given more challenging work, they do it." These schools appeared to be a likely group to translate ambition into practice.

The motives and practices for adding additional students to GT programs varied. In some schools administrators reported balancing class sizes more equitably by bringing more students into the program. Administrators noted they appreciated flexibility for adding students who were close in most areas of the criteria or were deemed especially promising. Some students accessed these classes because of considerable parental pressure on school leaders. Both teachers and administrators indicated this pressure correlated to the socio-economic status of the family.

> *"We do a lot of targeted things in ELA. Right now with the GT kids we're working on getting the commended ratings. Just because you're gifted doesn't mean you will perform well on the test. Our administration understands this, but they are pushing us for commended TAKS ratings."*

Most students in GT programs were reported to take Algebra 1 in 8th grade. As noted elsewhere, the significance of this opportunity was considerable. Students who access Algebra 1 in 8th grade had a clear path to taking calculus before graduating high school with the added incentive that if they attended college they would typically be eligible to begin their science sequence without math remediation. These students would also have the math necessary to take courses such as AP Chemistry or AP Physics earlier in their high school program. In high schools that weight grades, these additional AP courses, which are inaccessible to most students provide a considerable advantage in making the top 10%. When this group of students is removed from the drop out and "not college ready" statistics, those percentages balloon dramatically.[6] Few, if any, of the GT students required any special education services or learning (504) accommodations.

Perhaps the most dramatic differences for these students were the attitudes of their teachers toward what was taught and how instruction was designed.

---

[6] Note: for this study we were not able to access the longitudinal record back 6 years to determine the actual rate for EPISD. However, the general pattern in other districts guides this statement. GT students are a tiny percentage of high school dropouts as are their peers in advanced classes. Advanced classes include AP, IB, Dual Credit, and other similar programs that earn college credit. "Not college ready" applies to all students who must take non credit, remedial courses at college.

Consistently in study observations, GT classes engaged in presentations, hands on activities, and small group learning. While instances of these types of learning occurred in non-GT classes, they were more inconsistent or non-existent. GT teachers noted the pressure to find new ways of teaching familiar topics. For teachers that taught both GT and regular education classes, there was a considerable difference in the focus of the classes from "challenging learning" to "Getting ready for the TAKS."

> *"I teach differently than mainstream teachers. I can do more projects with my kids, and we also need more professional level products than regular education classrooms. I have smaller classrooms, and there's more flexibility in my classroom, so classroom management is different."*

Among their peers, GT teachers were consistently perceived as among the best teachers in the school. When asked informally what are the "must see" teachers, administrators and fellow teachers indicated the teachers of the GT programs. In observations, these teachers also seemed to be using more varied teaching styles and modalities than their peers. As a group the GT teachers tended to be relatively more experienced. The newest teachers were not teaching the GT classes.

Sharing the Rest

The first pieces of data associated with analyzing the relative equity of teaching and learning were class sizes and overall teacher loads as measured by number of students on their roster. Class sizes in middle school core areas, where every student took the same course, were expected to remain within a relatively small range of difference. "Teacher availability" reports produced by TEAMS showed these and a sample comparison of 7th grade ELA teachers follows (Table 4). Note: The total number of students for which each teacher was responsible is included.

Table 4. Sample Comparison English 7/Read 7 (names changed, real numbers)

Teacher	Total Number of Students	Class size	Range	Average Class Size
Fowler (6 Sections)	134	22,22,19,19,26,26	19-26	22.3
Jordan (6 Sections)	140	22,30,18, 22,30,18	18-30	23.3
Stockton* (6 Sections)	159*	24,24,29,29, 26*, 27*	24-29	26.5
Landry (G&T) (6 Sections)	76	10,10,10,12,12,12	10-12	11
Averages (includes all courses)	127	10-30		21
Averages for non G&T classes	144	18-30		24

*Extended class average included for missing two classes

Calculating average class size was often deceiving due to the dilution effect of a small section or a parallel course such as Grade 7, GT. Also, the likelihood that a regular education student would be in a class of 18-19 was just as likely as him/her being in a class of 29-30. The school could be more efficient by having half as many GT classes and using the staffing to reduce the overall student load for the remaining teachers or could choose to reduce staffing by .5 FTE; or, because this pattern is repeated over 3 grades, 1.5 FTE. This decision would not affect the desirable average class size of 24.

Table 4 above also reveals student access to the most respected teacher. Students who were not in the GT program had no chance of being in his/her class. Even in schools where the GT teacher taught in the core program, the chances were quite small of being in his/her class. No GT teacher taught an inclusion class. As was noted elsewhere in this study, GT teachers also received the highest quality professional development and got it more often than non-GT teachers.

Across schools, the system for assigning inclusion classes was inconsistent. Some schools used the content area teacher's name to build the full class roster. Others split classes and assigned students who were not special education to a content area teacher's roster and assigned special education students to the special education teacher's roster[7]. This discrepancy at the school level created an added layer of difficulty for both longitudinal data analysis and class distribution analysis. Specifically, the data system needed to account for special education students in the mainstream classroom to measure the effectiveness of inclusion classes, special education interventions, and supplemental instruction such as tutoring or Saturday school.

With regards to the online survey, 29% of the respondent staff believed, "No matter what happens at this school, many of our students will probably not graduate from high school." In follow up interviews a few themes emerged as for the reasons. Attitudes towards education, work ethic, and family support were consistent across all schools. However, no one suggested students could not graduate because of a lack of ability.

[7] This practice will also be examined in the section titled, *"What Can You Do?" – Addressing Student Level Challenges*.

"What Can You Do?" – Addressing Student Level Challenges

- Economically disadvantaged, at-risk and LEP students scored significantly lower on the math and reading TAKS; whereas students in GT programs scored significantly higher on the reading and math TAKS.
- From a data analysis perspective, we were unable to replicate the identification of GT students.
- Class populations in mainstream classes approached 50% special education in some classes.

Teachers and campus administrators described several student related challenges which impacted the delivery of instruction, and all but 6% of online survey respondents indicated their campuses actively sought ways to work with challenging students and student populations. Many of these challenges were related to student demographics and/or populations as follows: socioeconomic status (SES), the scheduling of special education students inclusion settings, and the impact of limited English proficiency (LEP) on TAKS scores and campus ratings.

Is Family Income Always A Predictor of Success or Failure?

Research consistently indicates there is strong correlation between socioeconomic status (SES) and student outcomes. Table 5 below outlines state and study school percent averages for SES, LEP, at risk, mobility and cumulative TAKS pass rates during 2010.

Table 5. TAKS Pass Rates

Middle School	Economically Disadvantaged	Limited English Proficiency	Classified as At Risk	Mobility Rate	TAKS Pass Rate, 2010*
Bassett	85.0%	23.5%	60.2%	31.3%	65%
Brown	32.9%	9.2%	34.2%	8.8%	86%
Canyon Hills	78.6%	16.1%	49.8%	17.5%	71%
Charles	75.0%	8.8%	43.5%	16.8%	79%
Lincoln	55.3%	14.0%	42.6%	12.3%	75%
Wiggs	76.1%	29.0%	58.6%	14.5%	74%
State Average	59.0%	16.9%	47.2%	--	77%

* Percent of students passing all tests, across all grade levels.

Results above did not necessarily align with state averages as it appeared some study schools performed at the same level as state averages or higher, in spite of larger economically disadvantaged student populations. A possible explanation for this discrepancy was that teachers and administrators associated with this study reported the academic culture of any campus could also influence student outcomes. In fact, teachers stated a genuinely caring school community could counter the potential negative impact of low SES. Teachers interviewed at four of

the five campuses reported a particularly strong sense of community that resulted in positive morale for teachers and students alike. Further, teachers at three of the five campuses reported strong administrative support and generous academic freedom that they stated afforded them the opportunity to better differentiate instruction to meet individual student needs.

Quantitative results indicated that among the student demographic variables examined, low SES students scored approximately 62 scale score points lower on the TAKS reading assessment than high SES students, while at-risk students scored 68 points lower than students not classified as at-risk. In addition, students identified as LEP scored an average of 112 points lower on TAKS reading than non-LEP students, while students in the GT program scored 46 points higher on the TAKS reading than students not participating in the GT program. Finally, students in the special education program scored 859 scale score points lower on TAKS reading than students not receiving special education services.

The analysis of TAKS math performance revealed low SES students scored 29 scaled score points lower than high SES students, though there was not a statistically significant difference between at-risk and non at-risk students on TAKS math. Similarly, there was not a statistically significant difference between LEP and non LEP students. Students participating in the GT program scored 47 scale score points higher. However, special education students scored 851 scale score points lower on TAKS math than students not identified as receiving special education services.

> *"The principal's leadership is a big factor of our success. When she came in she brought a new philosophy and brought the theme for the school: 'dignity and respect.' She brought in the weekly emails, she updates us on things, and she always keeps us informed. She doesn't talk about a lot of things at once. Instead, she targets a specific behavior, usually only one or two at a time. This approach is kind of manipulative, but in a good way. She balances her feedback with praise so you don't feel like you're constantly being corrected."*

The online survey supported interview data as 96% of respondents indicated their school's academic performance was based on the quality of instruction that children received. Further, 88% of respondents reported that administration in their school made a positive difference, and 94% reported that teachers made a positive difference.

Impact of Inclusion on Teacher Perceptions and Student-Teacher Ratios

Although 90% of online survey respondents stated that schools targeted specific areas for improvement, teachers expressed concern about current practices associated with working with special education populations in inclusion settings. Scheduling was reported as a contributing factor. In some schools a single classroom may have had 26 students present, but student names appeared on two different rosters: some students appeared on the roster of Teacher A, a regular

education teacher; and some students appeared on the roster with Teacher B, the special education teacher. These classes met in the classroom of the regular teacher, not the special education teacher.

The challenge with scheduling special education populations in this manner was two-fold. First, regular education teachers reported this arrangement resulted in the regular teacher often assuming more of the lead teacher responsibilities instead of feeling like they worked collaboratively with their special education colleagues as co-teachers. Instead, regular education teachers often reported their special education co-teachers worked one-on-one with students to assist as needed while the regular education teacher facilitated the lesson to the whole group. The effect was described as the regular education teacher is responsible for all of the planning and instruction of all students and the special education teacher was only partly responsible for the academic progress of some of the students. In contrast, special education teachers reported perceiving they had less academic freedom, were viewed more as a paraprofessional than a teacher, sometimes perceived they were not respected when classes were scheduled this way, and often indicated they had to shuffle between students in different classrooms during the same periods.

> Quote from a Special Education teacher – *"We have this inclusion model… where we are pulled between content areas and have to be in multiple classrooms at the same time. I personally think that because we don't have time to develop relationships with regular education teachers that they don't trust us to teach content. I can hear it in their voices when they talk to me."*

The second challenge associated with this scheduling practice impacts generating and monitoring student-teacher ratios. When student rosters are divided up between regular and special education teachers, even though those sections met at the same time and in the same classroom, the way teacher availability reports were generated gave an outsider an inaccurate picture of student-teacher ratios. Often campus administrators were the only ones who had a working understanding of the sections which met together to comprise one single class. The effect of this campus only based knowledge is that the data system cannot account for the link between academic teachers and specific groups of students for measuring effectiveness. Data has to be reprocessed by hand rather than accessing it technologically. The ongoing production of data reports, many times done by instructional coaches, wastes instructional time and frustrates staff.

Working with Limited English Proficiency Populations in the Borderland

El Paso is commonly referred to as "the Borderland" by locals and experiences a steady influx of Spanish speaking immigrants. Campus administrators and teachers consistently reported that students who were limited in their proficiency with the English language often struggled academically because reading, speaking, and understanding spoken language were fundamental skills needed to develop content mastery in all subject areas. No one school has an unequal share of these students, although some have predictably higher percentages than others. Although research shows that a minimum of seven years is needed before one is fluent in a second language, teachers indicated concern that LEP students are often only classified as LEP for the first two or three years that they are learning English. Teachers stated these students are prematurely labeled as regular education students during their third through their seventh years of acquiring English, and their results on TAKS assessments are included in determining whether or not campuses meet AYP requirements. This mandate creates unnecessary pressure on resources and students.

Instructional coaches reported identifying LEP needs by examining benchmark, TAKS, and mock TAKS results. Prior year data and mock TAKS data was also used for identifying mainstream students to target additional support. Although teachers indicated benchmarks were administered regularly throughout the academic year, mock TAKS were reported to be administered in January and February[8]. Following this, teachers, campus administrators, instructional coaches, retired teachers, paraprofessionals and/or tutors were reported to work with students who were identified to potentially need and/or benefit from TAKS tutoring.

[8]Unlike benchmarks, these mock TAKS administrations are not on the district calendar and are inconsistently implemented from school to school.

Extra! Extra! – Tutors, Saturday School, and Other Interventions

Regarding growth in TAKS math, results indicated Wiggs Middle School continued to decline in TAKS math performance over the past three years, while students at Lincoln Middle School remained rather static. In addition, students at H. E. Charles Middle school displayed marked improvement from year one to year two, but did not maintain the growth from year two to year three, where a decline in math performance was noted. As for reading, students at Lincoln, Charles, and Wiggs Middle Schools remained static in reading performance from year one to year two, but exhibited a slight decline from year two to year three. In contrast, students at Bassett Middle School exhibited a marked decline from year one to year two, but showed significant gain in reading performance from year two to year three. Based on the mixed results derived from the growth model analysis, the effectiveness of the interventions applied at each of the campuses was examined. Results indicated the extra time students spend in a course, such as TAKS reading and mathematics remediation, is not statistically significantly related to improved TAKS performance ($p > .05$). A plausible explanation for the findings, according to the assistant superintendent of middle schools, is that no standard curriculum is followed in these courses (Figures 9 and 10).

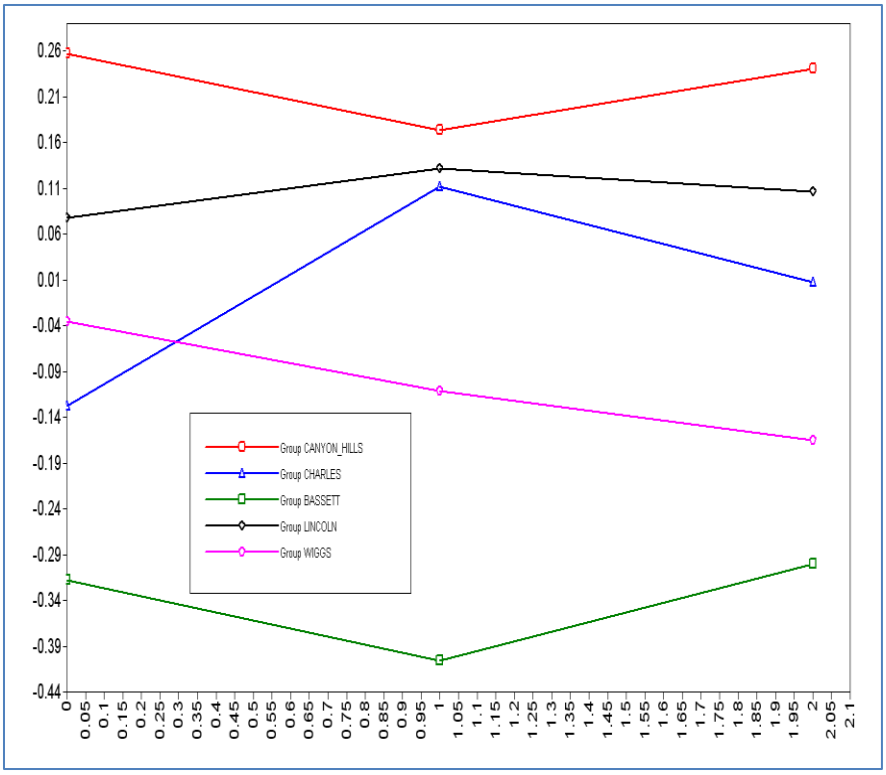

FIGURE 9. GROWTH IN TAKS MATH

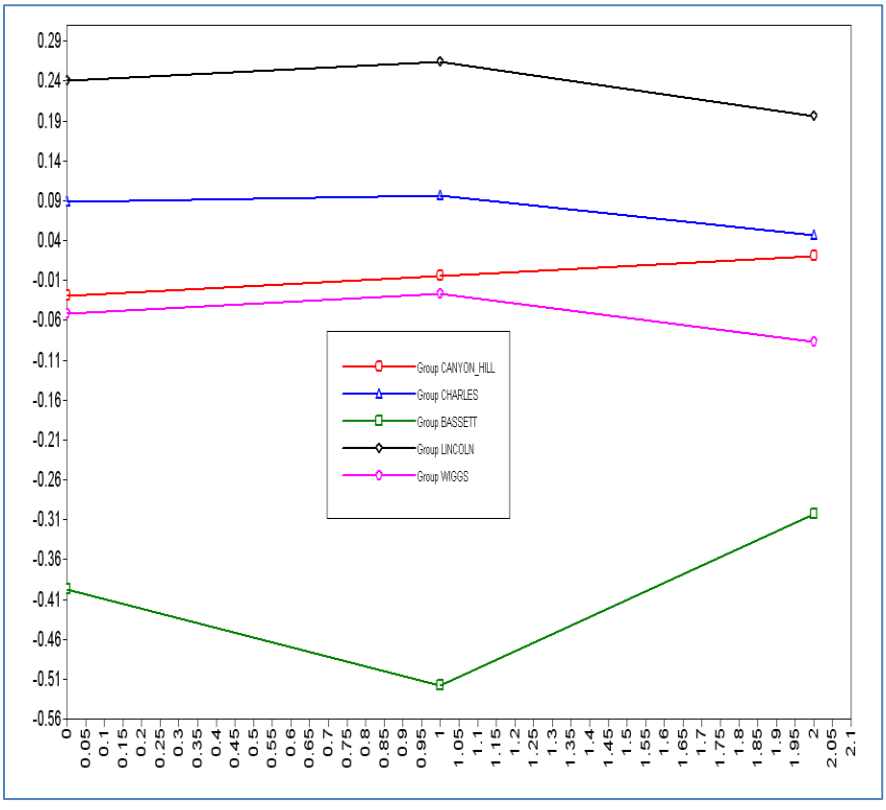

FIGURE 10. GROWTH IN TAKS READING

School leaders and teachers reported feeling significant pressure to have their students perform well on the TAKS. They indicated the district allows each campus to design intervention strategies that will be effective in improving the passing rates, and stated funding needed for this comes from campus budgets, supplemented where possible with Title 1 funds. The explicit goal of these programs was reported to be to improve student performance on TAKS. Programs and implementation models differ from campus to campus without a systematic method for evaluating their effectiveness or efficiency at the campus level or the district level.

Identifying Students

As previously mentioned, students who did not pass the TAKS in the previous year were automatically targeted for intervention by each campus. Additionally, students who performed poorly or unexpectedly low on benchmark and/or on assessments which utilized TAKS released items (referred to internally as "Mock TAKS"), were typically targeted. Students who showed a declining trend over the previous years - but who were not below the passing line - were not targeted for additional intervention beyond the school day. Identified students are typically scheduled for this intervention during an elective period. This practice keeps these students learning in these subject areas longer each day without the variety a typical middle school student experiences. It was not part of this study to interview students about the effect of this model.

Intervention Models

All students assigned to TAKS remediation classes took those classes in place of at least one of their electives. Examples of electives which might be given up for TAKS remediation included world language and keyboarding. Additionally, schools implemented three models (in varying configurations) of adding extra time to the regular school day for academic intervention. Some schools funded after school tutoring, while others offered voluntary "Saturday school" to reinforce learning. Sometimes teachers were paid for tutoring. However, other schools offered neither formal after school tutoring, or Saturday programs. Additionally, all schools encouraged their teaching staff to offer voluntary, supplementary support to students after school for all students.

During the regular school day, all schools hired additional tutors to either "push in" to classrooms for additional instructional support, or "pull out" students from their non-TAKS area class to meet individually or in small groups. These pull outs were done in coordination with the subject area teachers such as social studies, Spanish, or 7th Grade Science. Students missing their regularly scheduled class were responsible for making up any missed work. These interventions typically started in early winter. Training and supervising of these additional staff varied. One school actively deployed instructional coaches to train and monitor the tutors. In other schools, campus administration was responsible for the

"*Programs and implementation models differ from campus to campus without a systematic method for evaluating their effectiveness or efficiency at the campus level or the district level*."

> *If the tutors have the math knowledge and are willing to sit and follow your lesson, then they are very helpful. More students get one-on-one time that way. My tutor did his student teaching here last year, so he already knows how it goes. Our math coach also holds a session or two with the tutors and tells them what to do and what not to do. They are all certified as well… they all have a degree in something, but it might not be in math. My expectation is that they will reinforce what we're doing in my lesson. We want them to take the time to learn our methods and teach it to the kids the same way I would. Our tutors this year are making a "good" impact.*

supervision of their additional staff, but it was unclear how they monitored and evaluated the tutor's effectiveness beyond anecdotal feedback.

Teachers spoke positively about these interventions as time to build more constructive, responsive relationships with students. In these typically smaller settings, teachers reported they could better diagnose issues around learning on an individual basis and remediate effectively. Substantial doubt was expressed about the skills and abilities of tutors to recognize and remediate effectively.

Measuring Success

In part because of the unreliability of the data used to identify students or the intervention they received, no clear statistical result could be attributed to tutoring interventions. These interventions cannot be monitored effectively because participation – except for the tutoring during school hours – was voluntary for the most part. The predictably low level of student motivation for attending these sessions was a factor as well. Ironically, Saturday school had unexpected benefits as well. Some noted that attendance could be attribute to students having nowhere else to go and the chance to get something to eat.

Attendance and performance data could not be correlated because attendance was not consistently taken for these interventions. While a simple data collection system to monitor the effectiveness of these interventions could be conceived, much of the quality would not transfer. Overall, it was difficult to measure the effect of these programs other than through teacher and principal perceptions.

"A lot of kids who come to Saturday School come because they have to. You could do a lot more in Saturday school because you get more one-on-one time, and you could ask them to leave if they aren't focusing… which was nice."

"A lot of kids who come to Saturday School come because they have to. You could do a lot more in Saturday school because you get more one-on-one time, and you could ask them to leave if they aren't focusing… which was nice."

"It was decided that teachers would take care of their own tutoring of their own students. We didn't find that afterschool or Saturday school didn't work, but we didn't get the students we really needed. So we decided to use those funds in a different way."

In schools that abandoned Saturday programs, school leaders and teachers noted that students who were coming were not motivated to improve their performance, but rather by the rewards they received for attending. Principals reported after school tutoring had been made more difficult by the increased funding necessary for an after school transportation system and the logistics of getting students home. Consequently, all schools had some type of in school tutoring program. Schools tried to be more efficient in all of these programs by starting them later in the school year, closer to TAKS testing dates, and to a narrower group of students who were in need of remediation/intervention.

> *"The tutoring program is very important. We have a tutoring session after school. We had a lot of teachers involved with tutoring during PLC time to work on tutoring so that we could work on all the students… this was last year. Right now they are getting ready to start the process again to prepare for TAKS."*

Teachers indicated that in some cases, students struggled with content taught in September, though those students did not receive remediation for several months. The effect of the cumulative deficit and its implications on learning is a key feature of real time data systems. However, in the absence of such a data system, EPISD teachers noted feeling torn between choosing to keep pace with the curriculum or choosing to re-teach content and risk falling behind.

Teachers further stated EPISD's focus on correcting deficits after they are identified instead of preventing them beforehand compromised their relationship with students. Finally, it was noted that it was easier for a tutor who was skilled at introducing skills rather than re-teaching concepts in alternative methods. When viewed in the aggregate, no single implementation was shown to be more effective than another. Teacher and principal perceptions about these programs were also equally varied.

> *"Saturday school: we looked at the results of Saturday school students, we didn't see that it was helping in terms of data improvement on TAKS… but yet we were paying a teacher to meet with them, and also feeding the kids, etc. We then pulled in all the teachers who were involved to get their opinions. I discussed this with several different people and we decided not to do it anymore. We might have 2 teachers with only 10 students, and the kids didn't really improve. It wasn't cost effective."*

The most consistently favored model was to have tutors or other paraprofessionals present during the school day either be "pushed in" to classrooms with identified students or these tutors "pull out" students from non-TAKS area classrooms such as foreign language and keyboarding classrooms. No clear pattern or data emerged about the effect on student learning for students pulled out of their non-TAKS tested class.

DATA RENDERED USELESS IS POINTLESS TO KEEP

- EPISD spends roughly $300,000 annually to print district benchmarks exams
- Quality of benchmarks assessments are marginal at best
- The negative effects of these efforts on instructional time, teacher morale, student engagement, and administrative pressure are considerable

Student performance data was collected through the administration of a number of annual assessments. These assessments included common assessments and benchmarks as formative measures, and the Texas Assessment of Knowledge and Skills (TAKS), which will soon be phased out and replaced with the STARR, as its state mandated summative measure. Teachers indicated common assessments were administered weekly, whereas benchmarks were administered three times a year separately in each of the following content areas: mathematics, science, social studies and ELA/reading. Further, schools were afforded opportunities to administer Mock TAKS assessments and Mock STAAR assessments in the mid-winter, and were mandated by the state to administer spring TAKS tests in areas of social studies, science, ELA/reading, and mathematics. This section examines current practices and utilization of benchmarks, common assessments, and Mock TAKS assessments as associated with study schools.

PSYCHOMETRIC PROPERTIES OF BENCHMARKS

Although 91% of online survey respondents indicated teachers and administrators used assessment data to understand and provide more attention to the alignment of curriculum, the quality of the benchmark assessments among the middle schools examined was found to be marginal at best. From an efficiency perspective, this is disquieting considering that EPISD spends approximately $300,000.00 annually in printing costs to print benchmark assessments. Regarding psychometric properties, the assessments appeared to be limited in difficulty while providing no information about the ability of students with upper-level capability. The information curve depicted in the following figure is representative of the formative benchmark assessments in mathematics and reading administered in the middle schools during the timeframe of the study. Note the graphs below are from the 8th grade ELA benchmark assessment. To further explain, the information curve of a properly constructed assessment should be centered above zero. In the Figure 11 below, the information curve is centered at approximately -1.98.

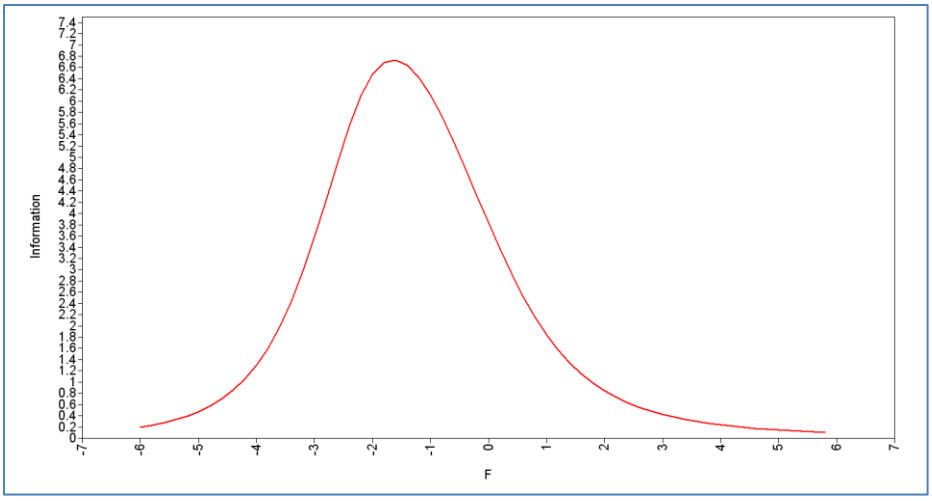

FIGURE 11. EIGHTH GRADE ELA BENCHMARK ASSESSMENT INFORMATION CURVE

Figure 12 below displays the item characteristic curve for the each question included in the assessment. For a properly constructed assessment, the questions should intersect at zero on the X-axis and .5 on the Y-axis. Difficulty goes from left to right, with less difficult questions on the left of zero (X-axis) and more difficult questions on the right of zero. The majority of questions displayed in the following figure are left of zero on the X-axis, indicating that a student with an average ability level has more than a 70% chance of obtaining the correct response. For a properly constructed assessment, a student with average ability (zero on the X-axis) should have a 50% chance (Y-axis) of obtaining the correct response. Based on the analysis, the assessments are not challenging students and the results obtained from the current assessments are of limited value.

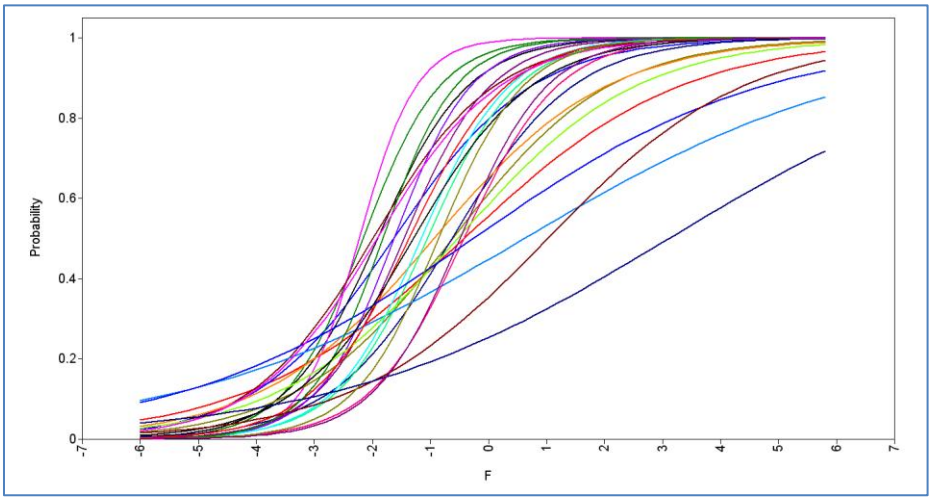

FIGURE 12. EIGHTH GRADE BENCHMARK BENCHMARK ASSESSMENT ITEM CHARACTERISTIC CURVE

Student and Teacher Perceptions of Formative Assessments

> *"I don't know if we use benchmarks the right way. We have a hard time with benchmarks because the kids know we can't use the benchmarks as part of our grades. The kids don't take them seriously so it's hard to use the data we get from them and apply what we learn to modify our curriculum."*

Teachers reported EPISD policy prohibits benchmark results being utilized in the calculation of student grades, although 91% of online survey respondents indicated assessment data is used to monitor student progress. As a result, teachers indicated student attitudes regarding their performance on benchmarks were poor. Teachers expressed concern that because students were not held accountable for their performance on any benchmark, students subsequently did not put for their best effort when taking these formative assessments. As a result, teachers generally did not perceive benchmark data as an accurate indicator of content mastery.

In contrast, common assessments were developed in response to current curriculum, expected to be administered weekly, and considered more relevant by teachers. Results from common assessments were reported to be incorporated into student grades, and teachers indicated students put forth more authentic effort when sitting for these formative assessments.

Alignment with Curriculum

Teachers and administrators stated that benchmark administration dates were announced by the district on relatively short notice, which made it difficult for teachers to develop instructional plans which either prepared for benchmark administrations or incorporated results of benchmark data. Teachers considered benchmark content to be largely separate from their curriculum, and reported test items were often unrelated to curriculum currently being taught. Teachers reported utilizing benchmark data predominantly by allowing students to review their graded benchmark tests during one class period, and then simply emphasizing specific test items students were generally weak on as lesson warm ups, ice breakers, or trivia type questions in following class periods. Teachers did not see the practical utility of current benchmark initiatives as valid, formative assessment tools.

> *"This year they mandated weekly common assessment and benchmarks so that makes it hard for us differentiate instruction. Because administration and instructional coaches are pushing for common assessments, it puts pressure on us as teachers. There hasn't been a good flow or scaffolding this year because of all the testing."*

Teachers viewed common assessments as more valid assessments, and indicated related test items were developed by PLC teams at each campus and that items

aligned more closely with current curriculum activities. Teachers reported common assessments as more reliable measures of curriculum than benchmarks. Although teachers preferred common assessment over benchmarks, they seemed to perceive common assessments as the lesser of two evils. Teachers also questioned the practical utility of being required to administer common assessments on a weekly basis.

Loss of Instructional Time

Teachers reported feeling overwhelmed by being responsible for administering multiple assessments throughout the school year. Table 6 below outlines the potential number of assessments that teachers and principals reported for administration during a typical school year. The loss of meaningful instructional time was a significant consequence of EPISD assessment initiatives.

Table 6. Number of Assessments During a Typical School Year

Assessment Type	Timing of Assessment	Number of Assessments/Year
Benchmarks	3 assessments per year in each of the following content areas: mathematics, science, social studies, ELA/reading	12
Common Assessments	Weekly for roughly 36 weeks, and across the following content areas: social studies, science, mathematics, ELA/reading	144
Mock TAKS/Mock STAAR	Varies, campuses may opt to administer or not administer Mock TAKS/Mock STAAR	Varies by campus
TAKS	Once per year in the following content areas (these vary by grade level): mathematics, science, social studies, ELA and/or reading	Up to 4 entire school days

IMPLICATIONS

This report is based on a comprehensive data extraction from EPISD[9], multiple multi layer regression analyses, teacher and administrative interviews, surveys, classroom observations, and document analysis. These implications provide a model for how this proof of concept has implications within and beyond EPISD.

EPISD WOULD BENEFIT FROM THE CREATION OF A SINGLE DATA WAREHOUSE THAT PROVIDES DYNAMIC, LONGITUDINAL SUPPORT TO TEACHERS AND ADMINISTRATORS.

- Educators today teach and work in complex environments in which the demands placed on them are exceedingly challenging. For more than two decades the education reform movement has focused on the improving instructional effectiveness. While this academic focus is not diminishing, there is a dilemma inherent in any attempt to increase instructional effectiveness. On the one hand, the education community now possess a greater knowledge than ever before of the factors that contribute to student learning. On the other hand, we continue to struggle to find ways to help teachers become aware of, understand, and apply this knowledge in schools and classrooms.
- The EDC pilot study clearly demonstrated that teachers need data in near real-time that is valid, reliable, and easy-to- understand to assist them in identifying not only which students are having difficulty, but why students are having difficulty *before they are unsuccessful* at meeting predetermined goals. The preventative model validated in this study shows great promise in providing actionable data in near real-time that can improve instructional effectiveness for individual students and teachers.

EPISD WOULD BENEFIT FROM THE ALIGNMENT OF FORMATIVE AND SUMMATIVE ASSESSMENTS WITH GRADES AND OTHER INTENDED OUTCOMES.

- It is recommended that EPISD examine their benchmarks in order to ensure that the exams are indeed providing a true measure of student performance, and that they re-evaluate the practical utility of administering common assessments on a weekly basis.
- Ideally, student performance is monitored and principals and teachers receive alerts when grades are not aligned with the curriculum or intended outcomes. The critical elements to this ideal include reliable and valid assessments. Options from published databases of items to self generated tests are viable, but the data system must support teachers' efforts to meaningfully use instructional time for purposeful student learning. The data system must assist educators with real time quality assurance of assessment items.

---

[9] The El Paso ISD Data Diagnostic and Planning Project Current State Report, which was conducted by the Texas High School Project during 2009 and 2010, can be found on the El Paso ISD website at www.episd.org.

- Implementing an innovative dashboard ensures that each grade earned by students is aligned with the intended outcomes of passing high stakes assessments, transitioning to the next grade, and graduating workforce and college ready. This dashboard should be adaptable to the technology skills of the users, but at a minimum provide easy to read monitoring.

EPISD WOULD BENEFIT FROM THE DEVELOPMENT OF AN INTEGRATED DATA SYSTEM WHICH LINKS TEACHER EVALUATION AND STUDENT PERFORMANCE.

- The current supervision model determines that principals must wait until the end of the school year (after the release of TAKS or STAAR results) to determine the instructional impact that a teacher had on her/ his students. Although this system of evaluating teacher effectiveness is common, waiting until the end of the year to determine the success of a teacher may be too late. The focus should be on prevention rather than intervention.
- Teacher effectiveness could be monitored in near real-time throughout the current school year using a specially derived longitudinal model in conjunction with an on-/off-track algorithm to determine each teacher's impact on keeping students on-track to pass state-mandated assessments at the end of the school year, transition to the next grade level, complete high school, and graduate workforce and college ready. This approach is based on well-designed, vertically linked assessments that monitor student growth and teacher effectiveness simultaneously in 6 week intervals.

EPISD WOULD BENEFIT FROM THE DEVELOPMENT OF AN INTEGRATED DATA SYSTEM THAT ALSO LINKS PROFESSIONAL DEVELOPMENT WITH DIFFERENTIATED, TEACHER SUPERVISION.

- From a preventative stance, principals can monitor each teacher's effectiveness score throughout the current school year from a dashboard that provides in-depth performance data about each teacher. This system has been tested in other states and proven successful. Based on near real-time monitoring of teacher performance results, principals can immediately determine areas of need and provide meaningful professional development to address identified areas when the problem is identified *during the school year*.
- EPISD can examine the content and quality of the staff development offered to ensure that the staff development is indeed aligned with teacher needs. Currently, the link between teacher supervision and professional development is ambiguous.

EPISD SHOULD CONSIDER A MORE EQUITABLE DISTRIBUTION OF EFFECTIVE TEACHERS ACROSS ALL CAMPUSES, BUILDINGS, AND GRADE LEVELS.

- Building on the prior implication, schools need to have a legitimate staff appraisal system that distinguishes performance among staff. This system can assist and help identify professional development needs as well as recognize outstanding performance. Additionally, teachers who are

particularly successful in specific high need areas – which can be as focused as a particular standard – can mentor and assist others.

- EPISD may want to consider examining student-teacher ratios by analyzing the number of students present in a classroom during a given period as well as by analyzing the class rosters of teachers. Once a more balanced environment exists, EPISD will be able to monitor students and teachers more equitably with no additional funding.
- Given the statistical ineffectiveness of the remedial TAKS courses, EPISD should consider altering those courses to achieve better use of resources and time to assist students on improving TAKS performance. Currently there is no consistent curriculum associated with those courses.

DISTRICTS SHOULD CONSIDER THE ANALYSIS OF EXPENDITURES IN THE FOLLOWING AREAS.

- To meet the impending budget challenges many administrators are reducing the number of teachers, cutting programs, and in some instances, closing schools to achieve financial efficiency, often with little or no *valid data* to assist in meeting the challenge. Although recent studies have compared schools and districts across the entire state (State Comptroller, 2011), this approach of using other districts as a comparison may not be the best method of determining areas in which to make budget cuts.
- Each district has unique characteristics that may preclude a fair comparison. The EDC efficiency models validated in the current study not only provided areas for potential savings targeted at individual campuses within a district, but showed how much each area should be reduced to achieve both instructional effectiveness and financial efficiency. In addition, the models identified peer campuses within the district that were operating efficiently. These peer campuses serve as reference points that inefficient campuses can refer to as they strive to determine how to become more efficient.
- Examine how Title 1 funds are expended in each of these areas among all middle schools receiving Title funds to ensure all campuses are operating at peak efficiency while providing the best educational program possible.
- We anticipate that efficient academic prevention (rather than intervention) measures will improve savings as well. The preliminary results also found correlations between building maintenance and health care cost of employees at those buildings.

Recommendations for Further Study

Research is currently underway to validate the effectiveness and efficiency models utilized in the current project in Socorro ISD, as well as to study transcript data from students in EPISD and Socorro ISD who attend El Paso Community College. The link between transcript data from public schools with student performance data in their first two years in college is being used to construct a post-secondary early warning system that will provide intervention points in middle and high school that can ensure all students graduate high school and are prepared for post-secondary challenges. In this study, both social and academic factors are considered. It is anticipated that the data derived from this study will be incorporated into each benchmark assessment and the on-/off-track algorithm will allow students and parents to determine if the student is on-track to graduate college ready as early as sixth grade. It is also part of the research design in these additional sites to survey students about effective and efficient practices.

Because of the limits of the existing data sets, no model was possible for tracking access to higher level classes, attendance, and success at high school from each of the middle schools. Over time, these trends could be factored in as success factors.

BIBLIOGRAPHY

Calkins, A., Guenther, W., Belfiore, G. & Lash, D. (2007). The turnaround challenge: Why America's best opportunity to dramatically improve student achievement lies in our worst-performing schools. Boston, MA: Mass Insight.

Charnes, A., Cooper W. and Rhodes E., (1978). Measuring the efficiency of decision making units," European Journal of Operation Research,2, pp.429-444.Coelli, T.J., Rao, D.S.P. and Battese G.E., (1998). An Introduction to Efficiency and Productivity Analysis, Kluwer Academic Publishers.

Cohen, J. (1988)., Statistical Power Analysis for the Behavioral Sciences (second ed.). Lawrence Eribaum Associates.

Creswell, J. W. & Miller, D. L. (2000). Determining validity in qualitative inquiry. Theory into Practice, 39(3), 124-13.

Darling-Hammond, L., & Berry, B. (2006). Highly qualified teachers for all. Educational Leadership, 14-20.

Gibbs, G.R. (April, 2006). Concordances and semi-automatic coding in qualitative analysis: Possibilities and barriers. Presentation to the Workshop on Bridging Quantitative and Qualitative Methods for Social Sciences Using Text Mining Techniques, University of Manchester.

Goldberg, P., & Proctor, K. (2000). Teacher voices: A survey on teacher recruitment and retention. Scholastic Teach Today and Tomorrow.

Guarino, C.M., Santibanez, L., & Daley, G.A. (2006). Teacher recruitment and retention: a review of the recent empirical literature. Review of Educational Research, 76(2), 173-208.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York: Routledge.

Ingersoll, R.M. (2001). Rejoinder: misunderstanding the problem of out-of-field teaching. Educational Researcher, 30, 21-22.

Johnson, S. (2000). Teaching's next generation. Education Week, 19(39).

Leithwood, K., Harris, A., & Strauss, T. (2010). Leading school turnaround: How successful leaders transform low performing schools. San Francisco, CA: Jossey-Bass.

Marzano, R. J. (2003). What works in schools: Translating research into action. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J. (2007). The art and science of teaching: A comprehensive framework for effective instruction. Alexandria, VA: American Association of School Administrators.

Marzano, R. J., Pickering, D.J. & Pollock, J. E. (2004). Classroom instruction that works: Researched based strategies for increasing student achievement. New York: Prentice Hall.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). School leadership that works: from research to results. Alexandria, Va.: Association for Supervision and Curriculum Development.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. Harvard Educational Review, 62(3), 279-300.

Maxwell, J.A. (2005). Qualitative research design: An interactive approach (2nd Edition). Thousand Oaks, CA: Sage Publications.

McCarthy, T.A. and Yaisawarng, S. "Technical efficiency in New Jersey school districts," in Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (Eds.), (1993). The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press: New York.

Morgan, DL. (1997). Focus groups as qualitative research. (2nd Edition) Thousand Oaks, CA: Sage Publications.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? Educational Evaluation and Policy Analyses, 26(3), 237-257.

Peterson, K. (2002). Effective teacher hiring: A guide to getting the best. Alexandria, VA: ASCD.

Stronge, J. H. (2007). Qualities of effective teachers. Alexandria, VA: American Association of School Administrators.

Temin, P. (2002). Teacher quality and the future of America. Cambridge, MA: National Bureau of Economic Research, Working Paper 8898.

Tough, P. (2009). Whatever it takes: Geoffrey Canada's quest to change Harlem and America. New York: Houghton Mifflin.

Zimmerman, F. (1991). The turnaround experience: Real-world lessons in revitalizing corporations. New York: Mcgraw-Hill.

APPENDIX A: BIOGRAPHIES

Jon H. Fleming

Jon H. Fleming is Chairman of the Board of Directors – The Texas Education Reform Foundation and Chairman and President of the Education Data Collaborative.

Fleming received his bachelors and masters degrees from Southern Methodist University and his doctorate from the School of Theology at Claremont (California) and the Claremont Graduate School. His Post Doctoral work was done at The University of Texas at Austin. His clinical work was done at The University of Texas Medical School at Houston – Department of Psychiatry and the Hermann Hospital. He was on the faculty in Psychiatry of the Medical School and served on the Committee on Human Experimentation and Protection of the Health Science Center as well as the Admissions Committee of the Medical School.

He was a senior planning officer of The University of Texas Health Science Center at Houston and Executive Vice-President and chief operating officer of Texas Womans University – Head of Health Science Education and Professor of Psychology. His academic career culminated in his presidency of Texas Wesleyan University where he also served as professor of psychology.

Dr. Fleming was appointed by Texas Governor Mark White to the Select Committee on Public Education chaired by Ross Perot. He chaired the sub-committee on "Educating the Child" (Education and Curriculum.) He was also vice chairman of the Sub-Committee on the Teaching Profession. Governor William P. Clements, Jr. appointed him to the Executive Committee of the Texas Criminal Justice Task Force. Fleming headed the Mental Health sub-committee of same and served as chairman of the Mental Health Advisory Committee to the Texas Department of Corrections. He also served as Chairman of the Advisory Committee to the Windham ISD of the Texas Department of Corrections. He served as co-Chair of the Steering Committee of the Texas eLearning Initiative.

Jimmy Byrd

Dr. Jimmy Byrd is one of the most respected psychometricians in Texas. He is an Associate Professor at the University of North Texas - Denton, in the Department of Educational Leadership, and a consultant with the Texas High School Project, Education Data Collaborative, and numerous districts across the state. In addition to his Texas work, he has done data analytics for the State of Arkansas Department of Education, Western Heights School District in Oklahoma, and is currently engaged in a study in New Mexico. Dr. Byrd has been a teacher, principal, and an assistant superintendent in Texas Public Schools in Burnet ISD and Richland Springs ISD. He holds a Ph. D. in Educational Leadership and a Ph. D. in Measurement and Statistics, both from Texas A&M University – College Station. He earned his master's degree in Educational Leadership at Tarleton State University, where he also completed his undergraduate degree in Agriculture Science. In addition to serving on various committees and advising doctoral students, Dr. Byrd serves on the Executive Board of the Texas Council of Professors of Educational Administration and on the Editorial Board of the School Leadership Review. His primary research interest is focused on developing longitudinal growth models and valid and reliable measures of student achievement. In addition, he has conducted

several studies relating to the use of data and the inadequacies of research methods commonly employed in educational research.

Willard R. Daggett

Dr. Willard R. Daggett is CEO of the International Center for Leadership in Education. He is known worldwide for his efforts to move education systems towards more rigorous and relevant skills and knowledge for all students. Dr. Daggett has assisted states, school districts and education ministries in other countries with their school reform initiatives. He is the creator of the Rigor/Relevance Framework, a practical planning and instructional tool that is used in schools throughout the United States. Dr. Daggett helps people look at education differently. He encourages his audiences to embrace what is best about our education system and to make the changes necessary to meet the needs of *all* students.

Douglas Silver

Douglas Silver has a wide range of experience in K-12 education. He taught high school English for eight years before becoming a building and the district level administrator. He served as a Connecticut state teaching evaluator and was part of several high school accreditation teams in the Northeast. Subsequently he worked at Scholastic Education leading research efforts, designing instructional models, and creating resources and materials for two dynamic new programs released in 2009, *Expert 21* and *Expert Space*. Currently he is the Director of Research at the Successful Practices Network.

Cynthia Williams

Cynthia Williams is an Assistant Professor of Professional Practice in the College of Education at Texas Christian University where her instructional focus and research interests lie in the area of initial teacher certification. She also consults and/or has consulted as a program evaluator for Fort Worth ISD, Dallas ISD, and the University of Texas at Dallas in the College of Engineering. She received a Bachelor of Science degree in psychology and business administration from Oklahoma State University, a Masters of Education degree in elementary and early childhood education from the University of Arkansas at Little Rock, and a Ph.D. from University of North Texas in educational research. She has taught in both private and public PK-12 schools in Arkansas and California. In addition to serving on various university and college committees, Dr. Williams advises undergraduate students and serves on master level examination committees. She also serves on the evaluation and assessment Committee for the North Texas P-16 Council and the statistical advisory panel for PEEQ, a current project examining teacher impact and effectiveness during the first three years post-certification.

Sylvia McMullen

Sylvia McMullen has 30 years experience working in education leadership, business development and marketing research. She is Project Director of the Education Data Collaborative and serves as consultant to companies, state departments, and school districts on data management systems. She also recently served as Project Director for the Nevada Race to the Top application. Other projects include the Texas High School Project Data Diagnostic on the Big 8 Urban Districts in Texas and the Arkansas Longitudinal Data System and Target Testing Project. Ms. McMullen is an experienced consultant in data systems and has worked with over 200 districts in the state of Texas. She is skilled in strategic systemic planning and educational

reform. She has a Masters of Educational Administration from Texas A&M and Doctorate of Jurisprudence from the University of Houston. She has served on numerous state organizations in Texas including the Texas Business Education Coalition and the Closing the Gap Initiative with the Texas Higher Education Coordinating Board.

APPENDIX B: DATA VARIABLES

This appendix lists the data variables used in this study.

ACADEMIC VARIABLES (OUTCOME VARIABLES)

- TAKS results for each year in reading, mathematics
- Benchmark / formative assessment results in each subject area from each year
- Other standardized test results (e.g., Stanford)
- Exam scores for all tests used to determine LEP status
- Teacher of record for mathematics and English by course, grade level, and year
- Course completion data (by service ID), including pass/fail and reason code, honors, dual credit, or remedial indicator, numeric end of semester grade, and teacher id for all summer courses AND credit recovery, as available.
- Curriculum Indicator (if used different curriculum during the timeframe of study)

STUDENT DEMOGRAPHIC VARIABLES (PREDICTOR VARIABLES)

- Gender
- Ethnicity
- Free or Reduced Price Lunch status
- Special Education status
- Gifted/talented status
- Disability status with primary and secondary disabilities and educational setting
- Limited English Proficient status
- English as Second Language status
- Migrant status
- Indication of student repeating grade or course by course, grade level, and year
- At-risk status
- Home language
- Student Entry Date into school (Month and year student entered school)
- Student Exit Date out of school (Month and year student left school)
- Grade enrolled by year
- Elementary Campus Attended (Name of elementary campus student attended)
- Middle School Campus Attended (Name of middle school campus student attended)
- Zip Code of student residence (Zip code of residence in each school year)
- Number of days absent each six weeks
- Number of days in membership each six weeks
- Number of discipline referrals each six weeks
- Type of discipline referrals each six weeks
- Student age in months

Teacher Variables (Predictor Variables)

- Teacher Experience (Years teaching experience)
- Teacher tenure at current campus (Years at current campus)
- Certification area(s) of teacher (Name of certification area)
- Teacher gender
- Teacher ethnicity
- Highest degree obtained by teacher
- Number of days teacher absent each year
- Number of staff development days in which teacher participated
- University or program where teacher received teaching certification (Name of University or Program)
- Campus where teacher is teaching by year (Name of Campus)

School-level Variables (Predictor Variables)

- Percent Mobility
- Total Campus Enrollment
- Student-teacher Ratio
- Percent Free or Reduced Price Lunch
- Percent Black Students
- Percent Hispanic Students
- Percent Limited English Proficient Students
- Percent At–risk Students
- Percent students graduating with the recommended high school diploma
- Percent teachers in first year of teaching

Campus Finance Variables (DEA Input Variables)

- Expenditure by Fund–General Fund (Fund 199)
- Expenditure by Fund–State Compensatory Education (Fund 185)
- Expenditure by Fund–Title 1- Part A (Fund 211)
- Expenditure by Function-Instruction Per Pupil
- Expenditure by Function-Guidance and Counseling Per Pupil
- Expenditure by Function-School Leadership Per Pupil
- Expenditure by Function-Instructional Resources and Media
- Expenditure by Function-Campus Maintenance and Operations
- Expenditure by Function-Total Operating Per Pupil
- Expenditure by Object–Payroll Costs (Object 6100)
- Expenditure by Object-Professional and Contracted Services (Object 6200)
- Expenditure by Object-Supplies and Materials (Object 6300)
- Expenditures by Object-Other Operating Expenses (Object 6400)
- Expenditure by Program-Regular Per Pupil
- Expenditure by Program-Bilingual Per Pupil
- Expenditure by Program-Gifted & Talented Per Pupil
- Expenditure by Program-Special Education Per Pupil
- Expenditure by Program-Compensatory Per Pupil

APPENDIX C: EL PASO MIDDLE SCHOOLS EFFICIENCY REPORT

In the following report, under the column labeled "Variable" TAKS is the Texas Assessment of Knowledge and Skills, Instruction refers to expenditure for the benefit of the classroom instructional program (Function 11), Instructional Resources refers to expenditures for resource centers, establishing and maintaining libraries and other major facilities dealing with educational resources and media (Function 12), School Leadership refers to expenditures associated with directing/managing a school campus (Function 23), Counseling refers to guidance and counseling Services such as assessing and testing students' abilities, aptitudes and interests; counseling students with respect to career and educational opportunities and helping them establish realistic goals (Function 31), and Plant Maintenance & Operation (Function 51) refers to expenditures to keep the physical plant and grounds open, clean, comfortable and in effective working condition and state of repair. Note the results are from fund 199 (General Revenue).

Under the column heading "Actual Expenditure", the figures associated with each function are the actual per-pupil expenditures expended in Fund 199 (General Fund) in the 2009-10 academic year. Under the column labeled "Goal Expenditure", the figures associated with each function are the per-pupil expenditures that should have been expended to obtain the actual percentage of students passing all TAKS in 2009-10. The difference between the Actual Expenditures and Goal Expenditures are measures of inefficiency or "Potential Improvement" and is the percent reduction in actual expenditures that would be necessary for each campus to achieve an efficiency score of 100%. More concretely, the percent reduction is a measure of inefficiency in each function examined. Line-item expenditures in each of these areas should be examined in detail to determine specific budget reductions are reallocation of funds to ensure maximum student performance in an efficient manner.

The dollar amounts under the column labeled "Potential Savings" are the estimated savings that could be realized by EPISD by either reducing the budget or reallocating funds to improve instruction. Realizing that drastic budget reductions or reallocation of funds in some campuses may not be feasible in one year, the potential savings serve as a goal that should be realized over a two year timeframe.

Lincoln Middle School **Efficiency Score: 100%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	75.00%	75.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3270.35	$3270.35	0.00%	
Instructional Resources	$54.78	$54.78	0.00%	
School Leadership	$502.01	$502.01	0.00%	
Counseling	$93.92	$93.92	0.00%	
Plant Maintenance & Operations	$153.74	$153.74	0.00%	

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Hornedo Middle School **Efficiency Score: 100%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	79.00%	79.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3039.56	$3039.56	0.00%	
Instructional Resources	$1.20	$1.20	0.00%	
School Leadership	$364.45	$364.45	0.00%	
Counseling	$193.63	$193.63	0.00%	
Plant Maintenance & Operations	$134.56	$134.56	0.00%	

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Brown Middle School **Efficiency Score: 100%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	86.00%	86.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$2869.90	$2869.90	0.00%	
Instructional Resources	$48.34	$48.34	0.00%	
School Leadership	$414.14	$414.14	0.00%	
Counseling	$143.95	$143.95	0.00%	
Plant Maintenance & Operations	$118.77	$118.77	0.00%	

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Magoffin Middle School **Efficiency Score: 89.06%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	79.72%	71.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3517.06	$2369.33	-32.63%	$1,057,059.33
Instructional Resources	$83.67	$39.91	-52.30%	$40,302.96
School Leadership	$471.83	$341.90	-27.54%	$119,665.53
Counseling	$150.75	$118.84	-21.17%	$29,389.11
Plant Maintenance & Operations	$110.10	$98.05	-10.94%	$11,098.05

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Richardson Middle School **Efficiency Score: 84.78%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	88.46%	75.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3491.71	$2776.27	-20.49%	$573,067.44
Instructional Resources	$76.35	$46.65	-38.90%	$23,789.70
School Leadership	$543.23	$411.35	-24.28%	$105,635.88
Counseling	$134.78	$114.27	-15.22%	$16,428.51
Plant Maintenance & Operations	$143.24	$121.45	-15.21%	$17,453.79

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Canyon Hills Middle School **Efficiency Score: 82.53%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	86.03%	71.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3344.27	$2581.30	-22.81%	$673,702.51
Instructional Resources	$72.16	$43.39	-39.87%	$25,403.91
School Leadership	$$461.42	$380.80	-17.47%	$71,187.46
Counseling	$133.42	$110.11	-17.47%	$20,582.73
Plant Maintenance & Operations	$159.53	$111.91	-29.85%	$42,048.46

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Morehead Middle School **Efficiency Score: 78.64%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	89.01%	70.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3482.24	$2738.56	-21.36%	$706,496.00
Instructional Resources	$69.46	$45.97	-33.82%	$22,315.50
School Leadership	$535.50	$410.97	-23.25%	$118,303.50
Counseling	$127.90	$100.58	-21.36%	$25,954.00
Plant Maintenance & Operations	$199.94	$122.98	-38.49%	$73,112.00

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Charles Middle School **Efficiency Score: 77.73%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	100%	79.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3553.03	$2761.63	-22.27%	$483,545.40
Instructional Resources	$109.70	$46.46	-57.65%	$38,639.64
School Leadership	$685.21	$403.43	-41.12%	$172,167.58
Counseling	$163.48	$127.07	-22.27%	$22,246.51
Plant Maintenance & Operations	$158.59	$117.29	-26.04%	$25,234.30

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Wiggs Middle School **Efficiency Score: 75.87%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	97.53%	74.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3254.75	$2469.45	-24.13%	$683,211.00
Instructional Resources	$78.20	$41.59	-46.82%	$31,850.70
School Leadership	$484.44	$356.35	-26.44%	$111,438.30
Counseling	$167.98	$123.86	-26.27%	$38,384.40
Plant Maintenance & Operations	$167.44	$102.20	-38.96%	$56,758.80

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Ross Middle School **Efficiency Score: 73.70%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	100%	74.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3425.65	$2473.80	-27.79%	$896,642.70
Instructional Resources	$62.04	$41.13	-33.70%	$19,697.22
School Leadership	$483.27	$356.18	-26.30%	$119,718.78
Counseling	$168.95	$124.52	-26.30%	$41,853.06
Plant Maintenance & Operations	$164.03	$102.47	-37.53%	$57,659.82

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Bassett Middle School **Efficiency Score: 71.05%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	91.49%	65.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3165.65	$2249.13	-28.95%	$816,619.32
Instructional Resources	$92.29	$27.96	-69.70%	$57,318.03
School Leadership	$436.10	$309.84	-28.95%	$112,497.66
Counseling	$178.07	$120.98	-32.06%	$50,867.19
Plant Maintenance & Operations	$214.91	$94.82	-55.88%	$101,288.88

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Terrace Hills Middle School **Efficiency Score: 65.15%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	100.00%	68.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3483.30	$2269.22	-34.85%	$755,157.76
Instructional Resources	$118.35	$38.22	-67.71%	$49,840.86
School Leadership	$660.82	$327.46	-50.45%	$207,349.92
Counseling	$194.78	$113.82	-41.56%	$50,357.12
Plant Maintenance & Operations	$236.79	$93.91	-60.34%	$88,871.36

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Guillen Middle School **Efficiency Score: 63.14%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	88.70%	56.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3578.79	$2005.51	-43.96%	$1,401,792.48
Instructional Resources	$88.14	$33.72	-61.74%	$48,488.22
School Leadership	$466.86	$294.76	-36.86%	$153,341.10
Counseling	$139.54	$88.10	-36.86%	$45,833.04
Plant Maintenance & Operations	$316.10	$86.27	-72.71%	$204,778.53

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Henderson Middle School **Efficiency Score: 60.92%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	90.28%	55.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3221.20	$1886.54	-41.43%	$1,233,225.84
Instructional Resources	$69.92	$31.75	-54.59%	$35,269.08
School Leadership	$450.15	$274.24	-39.08%	$162,540.84
Counseling	$147.65	$89.95	-39.08%	$53,314.80
Plant Maintenance & Operations	$162.45	$79.30	-51.18%	$76,830.60

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

Armendariz aka Cordova MS **Efficiency Score: 59.62%**

Variable	% Students that Should Have Passed All TAKS*	% Students that Passed All TAKS**		
TAKS	97.28%	58.00%		
	Actual Expenditure (per pupil)	Goal Expenditure (per pupil)	Potential Reduction (Measure of Inefficiency)	Potential Savings
Instruction	$3246.34	$1935.51	-40.38%	$1,036,866.53
Instructional Resources	$77.21	$32.60	-57.78%	$35,286.51
School Leadership	$547.01	$279.30	-48.94%	$211,758.61
Counseling	$183.16	$97.08	-47.00%	$68,089.28
Plant Maintenance & Operations	$218.63	$80.10	-63.36%	$109,577.23

*Based on actual expenditures

**The campus could have achieved the same results by reducing actual expenditures to goal expenditures.

APPENDIX D: STAFF SURVEY ANALYSIS

Successful Practices Network: Instrument E/E 2011 Teacher

Participants rated survey item based on a corresponding 4 choice scale that included 1 = Strongly Disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly Agree. The descriptive statistics including mean scores and standard deviations for each survey item on the SPN instrument are displayed in Table D-1. Average responses ranged from 1.66 (SD = .688) to 3.52 (SD = .541). The greatest amount of variation in responses was associated with item number six (SD = .964) regarding classroom management as a new teacher, while the least varied response among teachers was on item 43, which related to teachers using assessment data to understand and then provide more attention to the alignment of the intended, taught and tested curriculum. The overall mean response to this item was 3.31 (SD = .503).

To gain insight into the underlying structure of the SPN instrument, principal component analysis was conducted utilizing a Varimax orthogonal rotation. Based on the principal component analysis (PCA) and the results of the Parallel analysis (O'Connor, 2001), it was determined that five underlying constructs should be retained. As reported in Table D-2, construct one included fourteen items measuring the how school leaders improve and support the instructional program. Construct two included six items measuring teachers' use of assessments to understand and monitor student progress. Construct three included four items measuring instructional climate in the school. Construct four included six items measuring the importance of professional learning communities and collegiality. Construct five included four items measuring students' motivation to succeed. The five extracted constructs explained approximately 50% of the variance among total responses to the 48 item survey instrument. Reliability ranged from .70 (construct five) to .94 (construct one).

Nine survey items did not load into the first five constructs. These nine items clustered into an additional five constructs, but four of these remaining constructs were generated on the basis of only two items each, and the last construct was generated on the basis of only one item. Maxwell (2000) states that when constructs are formed with only two survey items, then the item carrying the largest pattern structure coefficient in that construct can be considered a solitary independent variable. Collectively, these last nine items accounted for 12% of the remaining survey variance. The resulting reliability estimates for each construct as measured by Cronbach's alpha are displayed in Table D-3.

Table D-1

Descriptive Measures of SPN Survey Items

No.	Item	N	Mean	SD	Min.	Max.
1.	ITEM	288	3.52	.541	2	4
2.	ITEM	287	3.34	.712	1	4
3.	ITEM	285	3.52	.547	2	4
4.	ITEM	270	3.26	.633	1	4
5.	ITEM	274	2.91	.798	1	4
6.	ITEM.	273	2.72	.964	1	4
7.	ITEM	249	2.75	.801	1	4
8.	ITEM	289	3.23	.763	1	4
9.	ITEM	291	3.37	.669	1	4
10.	ITEM	287	2.90	.669	1	4
11.	ITEM	288	3.16	.690	1	4
12.	ITEM	279	3.16	.747	1	4
13.	ITEM	285	3.27	.656	1	4
14.	ITEM	279	3.22	.702	1	4
15.	ITEM	282	2.89	.690	1	4
16.	ITEM	273	3.04	.839	1	4
17.	ITEM	272	3.29	.565	1	4
18.	*ITEM*	*279*	*3.24*	*.521*	*1*	4
19.	*ITEM*	*284*	*3.12*	*.672*	*1*	4
20.	*ITEM*	*284*	*3.52*	*.573*	*1*	4
21.	*ITEM*	*264*	*2.44*	*.801*	*1*	4
22.	*ITEM*	*278*	*1.78*	*.748*	*1*	4
23.	*ITEM*	*263*	*3.46*	*.550*	*1*	4
24.	*ITEM*	*278*	*3.06*	*.743*	*1*	4
25.	*ITEM*	*276*	*3.08*	*.685*	*1*	4
26.	*ITEM*	*272*	*2.39*	*.834*	*1*	4
27.	*ITEM*	*277*	*3.27*	*.546*	*1*	4
28.	ITEM	265	3.02	.868	1	4
29.	ITEM	221	2.87	.600	1	4
30.	ITEM	283	3.47	.560	1	4

No.	Item	N	Mean	SD	Min.	Max.
31.	ITEM	195	2.76	.830	1	4
32.	ITEM	247	2.08	.753	1	4
33.	ITEM	284	1.66	.688	1	4
34.	ITEM	277	3.47	.521	2	4
35.	ITEM	250	3.12	.832	1	4
36	ITEM	281	2.18	.905	1	4
37.	ITEM	281	3.15	.688	1	4
38.	ITEM	262	3.27	.551	1	4
39.	ITEM	284	3.24	.588	1	4
40	ITEM	257	2.75	.729	1	4
41	ITEM	264	3.13	.682	1	4
42.	ITEM	233	2.87	.910	1	4
43.	ITEM	274	3.31	.503	2	4
44.	ITEM	277	3.32	.533	2	4
45.	ITEM	271	3.15	.635	1	4
46.	ITEM	276	2.50	.868	1	4
47.	ITEM	270	2.86	.699	1	4
48.	ITEM	215	2.53	.906	1	4

Table D-2
Principal Component Analysis of Survey Items (Rotated Solution – Varimax Orthogonal Rotation)

Survey Item	Component										
	1	2	3	4	5	6	7	8	9	10	11
12.	.769										
13.	.760										
11.	.758										
14.	.728										
2.	.713										
25.	.703										
24.	.686										
8.	.644										
37.	.636										
27.	.572										
41.	.569										
9.	.568										
10.	.504										
38.	.490	.438									
43.		.727									
44.		.695									
39..		.572									
45.		.539									
18.		.450									
40.											
1.			.730								
3.			.727								
20.			.598								
4.			.552								
28.				.668							
16.				.616							
42.				.614							
19.				.572							
31.				.512							
17.				.448							
46.					.644						

Survey Item	Component										
	1	2	3	4	5	6	7	8	9	10	11
47.					.594						
15.					.520						
5.					.496						
34.						.716					
23.						.593					
30.						.538					
33.						-.514					
35.											
22.							-.667				
32.							-.474				
6.								.794			
7.								.781			
48.									.782		
21.									.654		
26.										.609	
29.										.536	
36.											.713
Eigenvalue 1.04	14.25	2.84	2.01	1.77	1.61	1.51	1.38	1.32	1.20	1.15	
% Variance 2.17	29.69	5.91	4.20	3.70	3.35	3.15	2.88	2.74	2.50	2.41	
Cum. Var. 62.72	29.69	35.61	39.80	43.51	46.87	50.01	52.89	55.64	58.14	60.55	

Table D-3
Reliability (Cronbach's Alpha) among Extracted Components

Component	Reliability
1	.94
2	.83
3	.77
4	.84
5	.70

EDUCATION DATA COLLABORATIVE
(a collaboration between Avid Stats, D2SC, SPN, and SchoolDude.com)

For further information about this report, contact:
Sylvia McMullen, Project Director
713-962-4520