

EDUCATORS' UNDERSTANDING OF VALUE-ADDED AND ITS IMPACT ON SCHOOL  
PERFORMANCE

A Dissertation

Presented to

The Faculty of the Education Department

Carson-Newman University

In Partial Fulfillment

Of the

Requirements for the Degree

Doctor of Education

Cory A. Smith

March 2017



Dissertation Approval

Student Name/ CNU ID: Cory Smith 0274964

Dissertation Title: Educators' Understanding of Value-Added and Its Impact on Student Performance

This dissertation has been approved and accepted by the faculty of the Education Department, Carson-Newman University, in partial fulfillment of the requirements for the degree, Doctor of Education.

Dissertation Committee:

Signatures: (Print or Type and Sign)

Dissertation Chair: Dr. Julia Price

*Julia J. Price*

Methodologist Member: Dr. Christopher Shon

*Chris Shon*

Content Member: Dr. Roy Miller

*Roy Miller*

Approved by the Dissertation Committee

Date: March 24, 2017

Copyright © 2017 by Cory A. Smith  
All Rights Reserved.

## Abstract

Confusion persists, and the debate continues regarding whether the current value-added models being widely used across the country are a reliable and valid measure of a teacher's competency. Educators are being Judged, but more importantly, do those working in the kindergarten through twelfth-grade public schools even understand the systems used to judge teacher and student performance? When reviewing the history of value-added models one could draw many conclusions. Some say the data is incorrectly used to discipline or fire teachers. While others argue, the data itself is fundamentally flawed. Specifically, when looking at local school district's data for both formative and summative assessments as it pertains to its use and potential use in teacher value-added scores, do the scores directly correlate to teacher effectiveness? As evidenced by many comments by teachers at local school board meetings during the 2014 and 2015 school years, many educators say they feel increased pressure to achieve results using a system they do not believe in or fully understand. At these school board meetings, teachers stated the Tennessee Value-Added Assessment System (TVAAS) measures created lower morale and many feel they are viewed only as a *number*. In this study, teachers understanding of the TVAAS, at schools with high TVAAS value-added scores will be examined to determine what themes led to the school's high performance on TVAAS value-added.

Keywords: *Value-Added (Tennessee specific measure); Value-added model (VALUE-ADDED MODEL, a generic system), Tennessee Value-added Assessment System (TVAAS); validity, effectiveness; evaluation accountability; educational policy; summative assessment, teacher moral; teacher motivation, professional development*

## Acknowledgments

First, I would like to thank Dr. Julia Price. You have the patience of a saint. Next, my other committee members, Dr. Christopher Shon and Dr. Roy Miller. Dr. Shon, thank you for leading me down a path that best met my intent and direction of study. Roy, you were the inspiration to return to graduate school and reach the next level. Finally, everyone who helped proofread this document. Your feedback was invaluable.

To my staff. The smartest, most dedicated educators one could ask for, without what you do for children daily any of my accomplishments are hollow. Remember only do about 50% of what central office tells you to!

To my parents. Whether separated by a few feet or 8,000 miles you have always been there and provided either an encouraging word or a pointed “conversation.” Somehow always knowing which one I needed at what time. So many kids fail today because what’s most important in life is learned at home, kindness, respect and perseverance.

To my wife, there aren't words to describe what I owe you. Where have the first 25 years gone? Through it all, you have been my stalwart. Without you, that skinny 20-year-old kid you married all those years ago would not have achieved what I have today. This degree is as much yours as mine.

## **Dedication**

To my kids, Hunter and Abigail. I feel the best measure of a person success is their children. Through the highs and lows, you are the best. I hope I have provided the guidance, motivation and been a strong enough role model that allows you to reach your dreams.

## Table of Contents

Chapter One: Purpose and Organization .....	1
Statement of the Problem.....	2
Purpose and Significance of Study .....	3
Theoretical Foundation .....	3
Research Question .....	4
Limitations and Delimitations.....	4
Assumptions and Definitions of Consensus .....	5
Chapter Two: Literature Review .....	7
Historical Context .....	7
Conceptual and Theoretical Framework.....	9
Professional Development .....	23
Summary .....	29
Chapter Three: Methods and Procedures.....	31
Definitions.....	31
Research Design.....	32
Researcher.....	36
Study Participants and Setting .....	39
Data Collection .....	39
Analytical Methods.....	40
Ethical Considerations .....	43
Summary .....	44
Chapter 4: Analysis of Data.....	45

Demographics .....	48
Data Analysis .....	49
Questions 1-10 .....	50
Member Check.....	66
Summary .....	68
Chapter 5: Conclusions, Implications, Recommendations .....	70
Conclusions.....	71
Implications.....	75
Recommendations.....	76
References.....	80
Appendices .....	93
A: Individual Data Feedback .....	94
B: Grade Level SMART Goal .....	99
C: Initial Interview Questions.....	101
D: Member Check Interview Questions .....	103

## List of Tables, Figures and Illustrations

Table 4.1 School Demographics .....	48
-------------------------------------	----

## **Chapter One**

### **Purpose and Organization**

The complex, dynamic and interactive nature of the learning process has consistently defied simple explanations and has placed significant conceptual and methodological challenges for researchers and practitioners who attempted to disentangle and isolate specific, direct effects on student achievement and growth. Two factors, in particular, that seem to be especially prohibitive are the dynamic, interactive nature of the learning process and the inevitable confounding of many of the formal and informal influences on the process (Holland, 2002).

The measurement of an educator's impact on student learning has always been a difficult subject to discuss and analyze. The passion involved when dealing with children can impede many a rational person. When introducing and explaining the background of value-added models and more specifically the Tennessee Value-Added Assessment System (TVAAS), it is important to remember the numbers analyzed are tied directly to a student or educator for whom judgments made concerning the individual's ability to do the job. While measuring performance has been used for decades in the private sector, it is a relatively new concept for educators. Student learning and development of academic proficiency are exceptionally complex problems influenced by several factors. "A few are personal characteristics (both cognitive and non-cognitive), physical and mental maturation, home environment, cultural sensitivities, institutional and informal community resources, and, of course, the formal process of schooling" (Koretz, 2001).

Over the last fourteen years, basically since the signing of the No Child Left Behind Act in 2002, there has been mounting pressure to measure teacher performance through the use of standardized testing. One of the major steps of data analysis, as related to student growth and performance, is value-added models of data analysis. The most common value-added model was developed by Dr. William L. Sanders at the University of Tennessee in the early 1980s and was first implemented in 1994 in the state of Tennessee. The goal was to develop a system irrelevant to a student's demographics, to measure student growth and generate information on whether a teacher is growing a student.

Dr. Sander's belief was that one can use a mixed statistical methodology to eliminate many of the previously cited barriers to measuring teacher impact on student growth. This method required three key components. One component was a testing process, which produced scales, that have a strong connection to the curriculum and which yield measurement that ranges above and below grade level. Secondly, the construction and ongoing expansion of a longitudinal database were noted. Finally, an analytical process that enabled a multivariate, longitudinal analysis to produce unbiased and efficient estimates of the desired effects was a method for determining the growth of students (Sanders, 2006).

### **Statement of the Problem**

Since the inception of value-added models, there has been much debate about its validity and merits. A review of professional literature of value-added models or Tennessee Value-Added Assessment System will find many articles both for, against and discussing the multitude of ways of measuring a teacher's performance. One area that has received little attention is the level of understanding of the educators who are measured by these various systems. With programs like merit pay, evaluation scores and overall job performance linked to value-added

model/TVAAS, one should be aware of how these system work. In this study, the researcher will investigate the following question. "Why do individual schools have high TVAAS value-added scores?"

### **Purpose and Significance of Study**

The purpose of the study was to determine why schools have higher TVAAS scores than others by using phenomenological research through interviewing educators to determine themes that may highlight aspects of an educator's background that influence school ratings. One theory was if educators who have a stronger understanding of TVAAS would then have better value-added scores leading to higher school scores. One theme that may appear was if these individuals participated in professional development programs, what level of influence on TVAAS value-added scores for the school did the professional development have? Did the professional development tie the assessed standards and curriculum together? These are some of the themes that were explored. Also, if this research revealed the relationship between teacher knowledge of TVAAS and a school's high value-added score or negative impact of teacher knowledge of TVAAS to student achievement, then this may lead to more intelligent conversations on better ways to measure school effectiveness.

### **Theoretical Foundation**

Professional development for teachers is now recognized as a vital component of educational practice to enhance the quality of teaching and learning in our schools. Consequently, an increased interest in research that identified features of effective professional learning has recently moved to the forefront of many educators' interest (Ingvarson, Meiers, & Beavis 2003). While many districts have promoted professional development for pedagogy and

best instructional practices minimal professional development has been implemented to educate staff on the test used to measure teachers' performance.

Employees in any organization need to feel valued. A 2012 study by the American Psychological Association reported the following comparisons of those who feel valued versus those who do not (APA, 2012). Feeling valued was a major factor in the employee's assessment of one's workplace. If an employee feels valued, they were more likely to report being motivated to do the very best for their employer. Whereas those individuals who do not feel valued were less likely to recommend a place of work to others and more likely to report wanting to find a new job in the future (APA, 2012). Add these components to the level of misinformation of value-added models, and it becomes clear why TVAAS can be a controversial issue.

### **Research Question**

“Why do individual schools have high TVAAS value-added scores?” The purpose of this study was to analyze influences an educator's understanding and knowledge of value-added models, specifically the Tennessee Value-Added Assessment System (TVAAS) has on the value added assessment scores for the schools in which the educators work.

### **Limitations and Delimitations**

The study is limited in population size to five schools and TVAAS value-added data currently archived with the state department of education. The demographics were limited to five schools in East Tennessee that utilized only the Tennessee Value-Added Assessment System as a formal measure of student growth. The study size had additional limitations based on the participation rate of educators at the schools chosen. Also, teacher's instructional abilities may or may not have fall into the same levels, however, by analyzing school-wide data, the variance in teacher's individual scores can be analyzed to determine trends in instructional approaches.

When considering value-added models, the multitude of research for possible consideration can be overwhelming. For this reason, in conjunction with the lack of research on teachers understanding of value-added models, the research question was tailored to a narrow scope to determine the level of a teachers' understanding of the process and how it may influence TVAAS scores for a school. Lastly, identifying additional impacts beyond a teacher's level of understanding of TVAAS in a school's performance was difficult using quantitative research. Therefore, a qualitative study using interviews was used to determine what experiences educators had with TVAAS that had any impact on student achievement.

### **Assumptions and Definition of Terms**

The overarching assumption was that many educators do not understand the system by which value-added and achievement data is determined. Also, there are four general assumptions about what Tennessee Value-added Assessment System (TVAAS) does. The first assumption was that TVAAS provides a valid model for judging student and school growth. Second, teachers use the same curriculum materials across same grade level classrooms and throughout the district. Third, TVAAS is a statistical methodology designed to evaluate the influence that school systems, schools, and individual teachers have on student learning. It was originally developed by Dr. William Sanders at the University of Tennessee. Finally, achievement measures a student's performance at one single point in time and is highly correlated with a student's demographics while comparing a student's performance to a standard. Outside of the generalized assumptions, the intent of TVAAS is to measure the growth of a student's academic profile across time and building on student data from one year to the next. For example, data across the years reflects only on test scores and is not related to a student's demographics. The

data should only compare a student's performance from one year to all the student's prior performance.

While value-added models were used to estimate or quantify how much of a positive or adverse effect individual teachers had on student learning during a given school year, it was also used to establish the fact value-added models in isolation cannot determine why a student may or may not have grown. It was determined educators know how to look deeper than just the numbers as to why trends occurred. Educators had to have sufficient knowledge of data mining to recognize important trends. However, determining the level of knowledge and skills leading to success on TVAAS was hard to measure through quantitative measures. Using a qualitative research process focused on educator's experiences resulted in the development of themes to determine the qualities of individuals who work in schools with high TVAAS value-added scores.

## **Chapter Two:**

### **Review of the Literature**

The term *value-added*, used in manufacturing to refer to the difference between the value of the output and the cost of the raw materials. In education, the term is used more loosely because value-added regarding changes in test scores is less tangible than value-added regarding some real currency (Hanushek, 1972). McCaffrey and Lockwood (2008) explained that while the origins of using value-added models to estimate teacher effects date back over 30 years (Hanushek, 1972; Murnane, 1975), interest in these methods grew precipitously following the publication of a technical report by Sanders and Rivers in 1996 (Braun, 2010).

### **Historical Context**

In response to the testing and accountability requirements of No Child Left Behind (NCLB), states and districts have been expanding testing programs and improving data systems. These actions have resulted in increasing reliance on student test score data for educational decision-making. Initially, NCLB focused on percentage proficient. A difficulty was that the percentage proficient does not give a complete view of student achievement. It did not provide data about the success or failure of students who are above or below the cutoff level. By contrast, value-added models did not take into account test score trajectories at all achievement levels. Furthermore, the percentage proficient was a problematic way to measure achievement gaps among subgroups of students. (Value-added workshop, 2010) The location of the proficiency cut score as related to the score distributions of the subgroups made a difference in the size of achievement gaps as measured by the percentage proficient (Holland, 2002). Additionally, the

exacerbating problem was looking at trends in achievement gaps with gaps in data not allowing for a complete picture of student growth.

The structural and functional features of the US educational system, learning environments present themselves as syndromes or amalgams rather than as additive clusters of independently accrued conditions (Rothstein, 2010). Students enrolled in Free and Reduced Lunch programs, for example, in addition to poor home environment, typically dealt with inadequate facilities, a less qualified teaching force, diminished curricula and conventional instructional methods, and explicit or implicit segregation along racial and ethnic lines. Consequently, these students consistently lag behind more privileged peers in academic achievement and progress (Brookhart, 1993). TVAAS developers have made a strong claim that system adequately accounts for all influences that potentially affect student learning. Considered issues like poor socioeconomic status and other factors are part of the statistical formulas, thereby allowing the sequestration of direct teacher outcomes on learning through employing the investigational design principle of *blocking*. Blocking requires using each student prior achievement as the only control or proxy for all such influences: "Each child, thought of as a blocking factor, enables the estimation of a school system, school, and teacher effects free of the socio-economic or other factors confounding what had historically been rendered unfair any attempts to compare districts and schools based on the inappropriate comparison of group means" (Sanders, Saxton, & Horn, 1994, p. 138).

One of the most rapidly advancing uses of test score data is value-added models which capitalizes on longitudinal data on individual students to inform decisions about the effectiveness of teachers, schools, or programs (Lockwood et al., 2006). Professionally, one would expect a teacher invested in best practice would continually reflect on the quality of job performance.

While the concept of self-reflection and examining best practices for teachers is not new, the current focus of using student summative achievement data for determining the level of teacher effectiveness is a more recent concept. These ideas were born from trying to find a way to measure if a perceived best practice is useful to move students with a variety of steps explored to include, administrative evaluation, peer evaluation, student survey feedback and review of annual student grades to name a few.

In 1971 Eric A. Hanushek stated there were three fundamental policy questions to be addressed: One, do teachers count? Second, are schools operated efficiently now? Third, what characteristics of teachers and classrooms are essential? In 1971, he wrote that educators the United States became concerned how educational services were being distributed, particularly along racial/ethnic lines. He proposed using a mathematical formula where one could determine the educational output of a student. Dr. Hanushek looked at family, peer, student's innate abilities, school inputs and a few other data points to try and determine the value output of a given student, teacher, school (Hanushek, 1971). "Surprisingly little is known about the ways in which schools and teachers affect education, largely as a result of past fixation on inputs to education rather than outputs" (Hanushek, 1971 pp. 280). When studying the work done by Dr. Hanushek, one could view it as the foundation for the value added system developed by Dr. William L. Sanders. Dr. Sanders research combined with the foundations laid by Dr. Hanushek used in combination with NCLB legislation were strong driving forces in education's current focus on value-added models.

### **Conceptual and Theoretical Framework**

In 2007, researchers asked the general public the following question: "In your opinion, which is the best way to measure the school's performance-by the percentage passing the test or

the improvement shown by the students longitudinally?" Eighty-two percent of respondents stated that the best way to measure school performance was to measure the gains posted by students longitudinally (Rose & Gallop, 2007). Measuring longitudinally would allow for the measure of *value* that the district, school, or teachers added to students' learning over time. The public's response was indicative of the overall trend in educational measurement and evaluation. The private sector understands the need for a quantitative measure of teacher effectiveness. However, before a discussion on value-added models can be had one has to know the difference between achievement and value added.

As measured by students' performance, achievement data viewed at a single point in time is compared to how well those students perform against a standard. Usually, student performance on state test is measured by how well students perform on state standards and the benchmark established for proficiency. District, school, and teacher success has been determined almost solely by the number of students who pass the state tests with minimal thought given to other possible measures (Battelle for Kids, 2014). The achievement data collected and measured by student performance on state tests results in teacher value-added scores. School performance is collective in nature, in that it was the result of the input of past teacher effectiveness, classroom peers, actions taken by administrators, and viewed as a cause for success or failure of the student (Harris and Sass, 2005).

Evidence-based theory embedded within the nexus of an ideological education reform effort in the elementary and secondary school environments has led K-12 public education to multipurpose the usage of student-generated standardized test scores (Price, 2014). Value-added is a statistical method used to measure teacher and school impact on student academic progress rates from year to year and is measured by how much gain or growth students make over time

(i.e., year-to-year, semester-to-semester) (Sanders, 2006). One way to think of academic progress regards a child's growth chart. A growth chart shows a child's height at age two, three, and so on. These data points can be plotted to display that child's physical growth over a period (Battelle for Kids, 2014). Similarly, if a student's math achievement level is measured annually using state growth data or nationally normed tests, the student's growth pattern in math can be plotted as it was for height.

With the work done by Dr. William L. Sanders, there has been a more determined push to measure teacher effectiveness through the use of "hard" data in the form of using student test data. Beginning in the early 1980's Dr. Sanders started developing and testing his methods and theories at the University of Tennessee, Knoxville. Initially, his work was not widely known and involved a small sample of Knox County public school students. As time progressed and a greater demand for accountability in schools Dr. Sander's work became increasingly influential as a measure of teacher accountability. Dr. Sanders used a mixed-model statistical methodology to develop his value-added model/value-added assessment measures.

Dr. Sanders' approach built on the work of Dr. Hanushek, and Sanders was able to isolate further particular factors that contribute to a student's academic success. Based on encouraging pilot studies with the value-added model conducted by Sanders and his colleagues during the 1980s, the Tennessee legislature has embraced the model as the methodology of choice to generate the desired evidence on the performance of students, teachers, schools, and school systems (Sanders, 2006). The legislation describes TVAAS as follows: A statistical system for educational outcome assessment which uses measures of student learning to enable the estimation of a teacher, school, and school district statistical distributions (Koretz, 2001). Also, the statistical system will use available and appropriate data as input to account for differences in

prior student attainment to plot the future growth of students. One concern expressed by some educators was the differences that the impact which the teacher, school, and school district have on the educational progress of students, may be estimated on a student attainment constant basis (Sanders, Saxton, & Horn, 1994; Sanders, 2006).

Currently, through some combination of a mathematical formula that looks at evaluation and test data has caused many teachers complain about evaluations reducing them to nothing more than a value-added score or number. However, what was Dr. Sanders' goal? Working with a fellow statistician, Dr. Robert McLean, they set out to eliminate many of the previously cited impediments to incorporating student achievement data in an educational outcome-based assessment system (Sanders, Saton, & Horn, 1994). The study showed the following: One, there are measurable differences among schools and teachers concerning effect on indicators of student learning. Two, estimates of school and teacher effects tended to be consistent from year to year. Three, teacher effects were not specific to a site; *gain* scores could not be predicted by only knowing the location of a school. Four, there was an adamant correlation between teacher effects as determined by the data and subjective evaluations by the supervisor. Finally, student gains were not related to the ability or achievement levels of the students when they entered the classroom.

Lastly, Dr. Sanders has said that scores for individual teachers should not be released publicly. Sanders further stated that releasing scores would be totally inappropriate and this is about trying to improve our schools, not embarrassing teachers. Sanders, (2006) "if teachers scores were made available, it would create chaos because most parents would be trying to get their kids into the same classroom." However, Sanders said identifying ineffective teachers is critical. The evidence is overwhelming that if any child catches two feeble teachers in a row

unless there is a significant intervention, that student never recovers from the negative experience, and that is something that as a society we cannot ignore (Hill, 2000).

While value-added models are widely used, currently 40 states and Washington, DC are using, piloting, or developing statewide growth or value-added models some individual states have gone further by developing state legislation and policies in addition to the federal requirements for teacher accountability (i.e., NCLB waivers and Race to the Top (RttT) grant requirements). A recent study indicated that 30 states and Washington, DC now have legislation or policies that require student achievement data be used to "significantly" inform the criteria for the evaluation of teacher effectiveness and subsequent decision-making efforts (Collins & Amrein-Beardsley, 2014). There are opponents of the system, with many educational bodies developing a more comprehensive look at teacher performance by including observation data, peer review, and student review, in addition to value-added data. There is no shortage of studies and controversy surrounding value-added models. Dr. Sanders also readily admits there can be issues with his value-added model.

In 2006, Dr. Sanders stated, "However, all value-added modeling efforts do not give equivalent results." A rejection of some of the simpler value-added approaches is appropriate because of severe biases and unreliable estimates that they provide (Sanders, 2006). When using value-added models as part of an accountability system, some minimum criteria must be required. To be able to partition student test score changes into pieces attributable to the district, the teacher and the student need data from other teachers within the district as well as the performance of the student with other teachers to be able to make comparisons. Thus, longitudinal data for each student is critical (Wainer, 2011).

When reviewing data to dampen the error of measurement associated with a single test score for an individual student, all assessment data over grades and subjects for each student must be used in the analysis. However, all students do not have the same quantity of test data. Disproportionately low scoring students have more missing longitudinal data than higher scoring students (Hill, 2000). Thus, any value-added model approach must be sophisticated enough to provide honest and reliable measures using all data for each student no matter how sparse or complete. Simple posttest minus pretest averages and simple regression approaches, which use only the previous year's score as a predictor variable, are examples of value-added model attempts that should not be used (Sanders, 2006). The use of student achievement data for decision-making is currently a focus of school and district reform efforts across the country (Strout, 2013).

In a local East Tennessee district, teachers regularly attend school board meetings to voice concerns about how data was used to measure teacher performance. However, among many educators there was confusion about what accountability TVAAS provides as a statistical measure of performance in the value-added model used by the district. A fifth grade teacher may be effective at teaching Roman history, but that is not a requirement measured by TVAAS. The impact which a teacher, school, or school district has on the progress or lack of achievement in the educational advancement or learning of a student is referred to as the *effect* of the teacher, school, or school district on the academic progress of students (Sanders, 2006). The statistical system shall have the capability of providing mixed model methodologies, which provide for best linear unbiased prediction for the teacher, school and school district effects on the educational progress of students (Corcoran, 2010). It must have the capability of adequately providing these estimates for the traditional classroom (one teacher teaching multiple subjects to

the same group of students), as well as the team-taught groups of students or other teaching situations, as appropriate. The metrics chosen to measure student learning must be linear scales covering the total range of topics involved in the approved curriculum to minimize ceiling and floor effects. These metrics should have a healthy relationship to the core curriculum for the appropriate grade level and subject (Education Improvement Act, 1992, §49-1-603).

Emphasis on data has grown as result of an increasing emphasis on using test scores to evaluate school performance, use that is central to the No Child Left Behind Act. A large part of the issues with value-added models is they are mathematically overwhelming to the lay person. To eliminate as many non-controllable variables as possible, statisticians have to add in as many controls as possible. The theory behind this is the more controlled/stable data points, the higher the correlation will be. An example of the calculations/formula used are:  $Y_{3\ 94} = b_{3\ 94} + u_{3\ 94} + e_{3\ 94}$  (1a):  $Y_{4\ 95} = b_{4\ 95} + u_{3\ 94} + u_{4\ 95} + e_{4\ 95}$  (1b):  $Y_{5\ 96} = b_{5\ 96} + u_{3\ 94} + u_{4\ 95} + u_{5\ 96} + e_{5\ 96}$ , etc. (Sanders, 2006). One can easily understand how educators are intimidated; even the most math-savvy teacher can see the complexity is daunting. Complicated formulas lead to people who are easily turned off by such measurements, label them unfairly, and end up intimidated by the whole process.

Adding to the confusion with overly simple blogs like efforts by Tim Childers (Childers, 2013), The Flaming Hoop of TVAAS where he gave a for too simplistic view of TVAAS. Also, laying out several falsehoods such as teachers are not allowed to see the expected growth for students and are not allowed to know who the *gap students* are. Also, he tries to add in multiple variables like tired, angry, or a student being absent for a few days before the test as possible reasons for poor teacher results on value-added models. Bloggers like Mr. Childers conveniently

leave out the fact that larger the data set, the more reliable the data becomes, as the pool of outliers will be significantly smaller than those who did not have issues.

Another often overlooked item is the longitudinal data set of TVAAS, that is the period for value-added scores averaged over a three-year average. At minimal, using longitudinal data in TVAAS one can only compare a child's value-added scores to expected growth from previous years scores, as compared to his peers. When using TVAAS to determine an individual teacher effect score, the data set must include three years of data. Any less than three data points and there is limited to no correlation. The lack of data points in determining effect scores is clearly indicated in a study by Sean Corcoran, "With one year of data, it is impossible to separate teacher effects from other factors affecting the performance of students within a classroom such as a disruptive classmate" (Corcoran, 2010).

While an unusually high set of test-score gains is suggestive of an effective teacher, one can only know if these gains are systematic after additional years of observations (Cruickshank & Haefele, 1990). These extra years are of particular importance when scores are "noisy," that is, when achievement gains are not well explained by variables accounted for in the statistical model. The longer time frames are, the more precise the value-added score estimates. These become more precise with each additional year of data (Corcoran, 2010). Additionally, one must realize what a value-added score does and doesn't tell us. Value-added models simply report how much a student did or did not grow. These do not measure why the student did or didn't experience success. Additionally, in a discussion of research-based indicators of effective teaching, Cruickshank and Haefele (1990) stated, "An enormous underlying problem with teacher evaluation relates to a lack of agreement about what constitutes effective teaching" (Cruickshank & Haefele, 1990, p. 34).

Value-added scores should be seen as a jumping off point for further analysis of a teacher's practice. Once a teacher is showing positive value-added scores, it is the administration's job to conduct a thorough review of the teacher's practice to find out what is causing the increase in scores. To the extent that principal judgments focus on identifying the best and worst teachers, for example, to determine bonuses and teacher dismissal, the evidence presented here suggests that such evaluations would help promote student achievement. Moreover, principal assessment has the potential to mitigate some of the concerns regarding strategic behavior on the part of teachers to improve test scores without increasing actual knowledge (Jacob & Levitt, 2003).

Several major newspapers have done in-depth articles on value-added models. In the *Washington Post*, May 9, 2011, reporter Valerie Strauss does a solid review of value-added models. However, she as many other detractors of systems based in value-added models, focused on slightly more intangible evidence of why a value-added model may be flawed and may not be the best measure of teacher effectiveness. She refers to Campbell's law, "The more any quantitative social indicator used for social decision-making, the more subject it will be to corruption pressures and the more likely it will be to distort and corrupt the social processes it is intended to measure" (Campbell, 1976). In other words, people will be tempted to cheat to get the reward. Cheating is no more evident than in the case recently concluded in Atlanta where over 20 educators were sentenced to jail for roles in inflating student scores to collect bonuses. The superintendent alone claimed over \$500,000 in bonuses and incentives (AJC, 2015). However, with the ability to cheat eliminated, there is a high correlation in state developed value-added models. A Pittsburgh study found that value-added model estimates based on local assessments designed to align with the curriculum (though not developed by psychometric experts) are

reliable enough to reveal statistically significant distinctions among teachers (Johnson et al., 2012).

Value-added models have become even more controversial with their increased use for incentive pay systems. In support of individual educators who are disputing lost bonuses because of low TVAAS scores, the Tennessee Education Association (TEA), has filed three lawsuits in Tennessee in 2010, 2014, and 2015 (TEA, 2015). While there may be merit to the cases because of the loss of incentive pay, mainly due to a small sample size of students or teachers having to accept school-wide scores because they have no classes with students who would have individual teacher measures the case does not deal with what TVAAS was designed to do, measure student growth. Classes like Art, Music, and Physical Education are all examples of teachers who benefit from high school scores. While many teachers claim value-added models are unfair, deeper consideration of why teacher unions fight these systems so much is warranted. Historically, from the National Education Association (NEA) perspective, the weaker the evaluation system, the better for teacher rights. Unions and teacher opposed to standardized testing will often claim we are teaching to the test (Overman, 2012).

A widely cited conclusion from the Sanders and Rivers study (2009) states: "Based on these results, students benefiting from regular yearly assignment to more effective teachers (even if by chance) have an extreme advantage regarding attaining higher levels of achievement" (Sanders & Rivers, 2009, p. 7). Sanders and Rivers examined the consequences for a student's performance and growth with teacher assignments over a three-year period. The researchers reached the conclusion that students who had ineffective teachers had a dramatic difference in performance compared to students who had effective teachers. However, these results can only

be taken to be insightful if we choose to ignore the definition of teacher effectiveness regarding students' performance gains as defined by TVAAS.

Accompanying state assessment programs, nearly all states have developed common school report cards, based on national and state standards, for disseminating information to the public about school quality (Deslandes et al., 2009). Reports cards are supposed to be a period report of student progress throughout the year based on homework, test scores, and student participation in class. Test are often weighted more heavily than the other measures. Grades have long been identified by those in the measurement community as prime examples of unreliable measurement (Brookhart, 1993; Stiggins, Frisbie, & Griswold, 1989). What one teacher considers in determining students' grades may differ substantially from the criteria used by other teachers (Cizek, Fitzgerald, & Rachor, 1996; McMillan, Workman, & Myran, 1999). Even in schools where established grading policies offer guidelines for assigning grades, significant variation remains in individual teachers' grading practices (Brookhart, 1993). Because of unique grading adaptations made by nearly every teacher, this variation in grading is even wider for students with disabilities and English Language Learners (ELLs) (Polloway, et al., 1994). Most statewide testing programs rely primarily on conventional multiple-choice tests because of low cost, ease and consistency of scoring, and a mature industry of companies offering a comprehensive menu of services for administering, processing, storing, analyzing, and reporting test results, ensure the privileged status of multiple-choice tests (Brookhart, 1993).

With some options, it is easy it is easy to see why policymakers mandate standardized testing; they are easy to administer and to score. The results translate into tidy percentages that impress parents and the media with the perceived comparability they offer between students,

teachers, schools, districts, and states, and even nations. They seem like a concrete way to evaluate teachers, schools, and districts (Herndon, 1985).

TVAAS is deeper, looking from one year to the next on student progress instead of quarterly, as report cards do. Also, TVAAS tries to be objective as possible, removing the human element of grade inflation. The data spread for value-added scores at one high performing school correlates well with other teacher measures of student progress, such as report cards and unit/module tests. While there is a growing consensus that teacher quality is important and current evaluation systems are inadequate, many have expressed concerns over the use of value-added models in high-stakes personnel decisions. There is a consensus that effective teaching matters. It may be the single most important school-based factor in improving student achievement (Darling- Hammond, 2000; Wright, Horn, & Sanders, 1997).

Measuring teacher effectiveness has remained elusive in part because of the ongoing debate about what an effective teacher is and does. These concerns are often grounded in the statistical imprecision and possible susceptibility to bias of the value-added (Briggs & Dominique, 2011; Corcoran & Goldhaber, 2013; Harris, 2011). Because of the lack of random assignment of students to teachers, value-added model become plausibly biased by the presence of other unmeasured students, class, or school influences on achievement (Rothstein, 2010). Moreover, value-added model is imprecise, with a substantial proportion of the variation in achievement across classrooms attributable to the student and classroom-level noise (McCaffrey et al., 2009; Schochet & Chiang, 2013). While research on these issues is ongoing, the prevailing view appears to hold that these limitations are not significant barriers to the use of value-added models in evaluating teacher performance, if done with appropriate caution (Glazerman et al., 2010/2011).

The Tennessee Value-Added Assessment System (TVAAS) is the most widely known application of value-added models in the United States, and efforts to extend or replicate this model are currently underway in other states and school districts (McCaffrey & Hamilton, 2007). After the initial work done by Dr. Sanders, many researchers have added other measures to increase the validity of teacher evaluations. While valuable for large-scale studies, some argue serious flaws exist in value-added models for evaluating individual teachers, and that rigorous, ongoing assessment by teaching experts serves everyone better (Sanders & Horn, 1994). Indeed, reviews by the National Research Council, the RAND Corp., and the Educational Testing Service have all concluded that value-added estimates of teacher effectiveness should not be used to make high-stakes decisions about teachers (Darling-Hammond, 2012). By focusing on outcomes of value-added model scores rather than the test processes educators and schools can freely use whatever methods prove practical in achieving student academic progress. TVAAS does not assume a perfect teacher or the best way to teach. Rather, the assumption is that effective teaching, in whatever form will lead to student gains" (Koretz, 2001).

When statistical estimates become a part of the procedure for summative evaluation of teachers, fairness is a key consideration (Goldhaber & Hansen, 2010). In the TVAAS model, teacher effects are shrunken estimates when not enough student data is available; a teacher is assumed to perform at the level his or her school system mean. The fewer students a teacher has, the stronger the pull toward the overall system mean. "A critical consequence is that it is nearly impossible for individual teachers with small quantities of student data to have estimates measurably different from the system means" (Sanders & Horn, 1994). Also, not only do teachers typically receive scant feedback on past performance in raising test scores, the

information they usually receive on the average test scores or proficiency of students can be misleading or demoralizing (Goldhaber & Hansen, 2010).

It is not easy to tease out all the issues surrounding value-added. In one survey almost half (46%) of a sample of Houston Independent School District (HISD) teachers who moved to different grade levels reported switching value-added ranks after the move, from "ineffective" to "effective" or vice versa (Collins, 2014). This is problematic as the SAS® EVAAS® system is purported to measure the teacher effectiveness construct consistently and with validity. Dr. William L. Sanders, the developer of the SAS® EVAAS®, claims that teachers who move from one environment to another, even if radically different, continue to do just as well (LeClaire, 2011). Furthermore, over half (55%) of the same sample of HISD teachers noted that SAS® EVAAS® reports did not match the supervisors' observational/performance scores (Collins, 2014).

Studies of other state proposals suggest that instruction tends to be resilient to externally mandated policies. The educational initiatives, when compared to instructional actions, are often far removed from the home of the effort, in addition to generally being poorly supervised and evaluated (Meyer & Rowan, 1978; Rowan, 1982; Weick, 1976). Furthermore, the incentive configurations essential to guarantee fidelity of enactment are usually missing (Elmore, 1996). This lack of correlation with policy and instruction makes any universal reform demanding (Elmore & Fuhrman, 1995).

Lastly, some suggest that supervisors are skewing observational scores to match teacher SAS® EVAAS® scores, given external pressures to do so (Collins, 2014). Such practices have been shown to occur elsewhere with the Tennessee Value-Added Assessment System (TVAAS) from which the SAS® EVAAS® was derived (Garland, 2012). In New York, if teachers have

two years of low value-added scores, the teachers are to be rated ineffective overall and terminated, regardless of what other measures (e.g., supervisor evaluation scores) indicate or disclose (Ravitch, 2012). When using less objective methods like supervisor evaluations and other methods to evaluate teachers, it seems that determining teacher effectiveness using value-added models is starting to outperform other indicators in capturing what it means to be an effective teacher. Such practices also negate the field standards developed by national associations on educational measurement and testing.

Knowing what the value-added models in place were designed to do may help eliminate some of the skepticism prevalent in today's educational environment. John Easton described a phenomenon in Chicago: essentially the "lack of internal capacity to use information profitably," but he nonetheless believes that value-added models can be utilized for research and evaluation and eventually to identify good school-wide and classroom-based teaching practices (Braun, 2010).

### **Professional Development**

While many teachers acknowledge facing issues such as student apathy, lack of parental support, high student mobility, cultural differences, negative attitudes, and absenteeism that may impact a student's performance; many cannot elaborate on how the system used to measure overall effectiveness works. Can knowledge of the TVAAS system overcome at least some of these concerns which the TVAAS system claims it factors out? While looking at educator understanding of TVASS, other areas should be considered. One area to consider is the practical strategies high performing teachers use. For example, regularly identifying the standards for students, telling students what they were supposed to learn and why, and the use of constant instruction to inform instruction. When studying how an educator's knowledge can affect

performance, one needs to consider all challenges faced in helping students succeed such as incomplete professional development, lack of resources, or poor leadership. However, by educators having a deeper understanding of the systems used to measure results, may improve results.

The fundamental purpose of examining and reflecting on students' needs in Professional Learning Communities is to cultivate ongoing expert learning for individual teachers. Additionally, to steadily expand best practices, where "teaching is a nuanced dance in which teachers integrate knowledge and pedagogical content knowledge, to be responsive to students' needs" (Timperley, 2011, p. 16). Understanding the role professional development in an improving a teacher's practice is crucial.

Teachers not only need to know how to teach but what to teach. Equally important and not as frequently discussed teachers need to know how the assessments of practice and instruction work. Studies show with professional development teacher's pedagogy improves individual effectiveness and is composed of two distinct components. One, the personal competence of one's implementation of teaching practice relates to proficiency and ability to influence student learning. Additionally, directly tied to the personal level of influence personal competence. Competence being a teacher's perception of his or her ability to operate at a high degree of proficiency in a particular domain and personal level of influence the concept of an educator's view of how much impact they have on student learning (Hoy & Woolfolk, 1993).

Specifically, professional development can change teachers' beliefs about individual and collective efficacy. Both types of effectiveness are critical to educators' persistence, drive, and success (Zimmerman, 1995). Highly associated with teacher motivation is Individual Efficacy. Which, in turn, affects student achievement (Bandura, 1993/1997). Teachers with a strong sense

of personal efficacy tend to spend more time planning, designing, and organizing what they teach. They are open to new ideas, willing to try new strategies, set high goals, and persist through setbacks and times of change (Goddard & Woolfolk, 2000).

Teachers continue to have multiple demands placed on them. To understand the governmental, institutional, and local requirements on teaching and new pedagogical practices is one in which teachers cognitively and emotionally attempt "to fit new information into existing knowledge and beliefs" (Ketelaar, Beijjaard, Boshuizen, & Den-Brok, 2012, p. 274). Previous understanding defines teachers' construct of new policies and practices, and it affects the way in which they will integrate initiatives into teaching (Allen & Penuel, 2015; Cohen & Hill, 2001; Jennings, 1996). Addition, as teachers make sense of new policies and practices, they do so in the context of their classrooms, school, and the district as well as about peers and students (Coburn, 2001; Ketelaar et al., 2012). Overall, own knowledge, level of experience, the context of the environment, and collaborations with administrators and colleagues improve classroom performance. Student performance affects the ways teachers make sense of and apply new policies/reforms and PD in the classrooms (Allen & Penuel, 2015; Coburn, 2001; Ketelaar et al., 2012; Lipsky, 1980). In total, policymakers who implement instructional policy must take into account the impact of all factors on a teacher's ability to interpret the policies implemented by the politicians that direct teachers to change instruction practice. When teachers understand the *why* of a change and the how the change works, results should improve.

NCLB and the development of value-added models have pushed the pace in education policy changes and directly affected the speed of pedagogical change for teachers. Would it alter the goal of an organization if shifting the purpose of professional learning was self-serving for the school or district? Potentially, the outcomes could lead to more individuality and sovereignty

for teachers in an organization. Sovereignty, or more individual freedom for instruction, in turn, may create an increased sense of ownership over the instructional process for staff. If a group can create knowledge for itself then the team is building a shared sense of meaning and developing the culture within the organization (Cranton, 2006; Yorks & Marsick, 2016). Furthermore, with increased data from value-added models "transformative learning becomes an expansion of consciousness that is collective as well as individual" (Cranton, 2006 p. 48).

This expansion of outcomes would hopefully be a collaborative environment created as educators work together to examine and question pedagogy, reflect on data, and share personal stories with others (Mitchell & Sackney, 2009). Within the culture of collaborative learning, teachers are consistently encouraged and supported to experiment with innovations and have a sense of professional entrepreneurship (Hargreaves & Fink, 2006). At issue is the value-added models born from NCLB have placed control measures on schools not developed locally. Instead, many of these directives have relied on external control measures potentially reducing the traditional changes drawn up by schools relying on purposes, professional socialization, collegiality, and natural interdependence. Furthermore, when staff can have shared expertise, this allows pedagogy to grow. When teachers can develop knowledge together, as a collective, it strengthens the group as a whole and elevates the members to an equal level of expertise, rather than a hierarchal structure of specialization (Sergiovanni, 1994). When teachers collaborate, it leads to a greater standard of trust and buy-in. One of the key factors is to get teachers talking with others about best practices, challenges, and successes, as well as other major educational trends. Indeed, "teachers cannot readily engage in cycles of inquiry and knowledge-building when they feel criticized or put down for not being good enough" (Timperley, 2011 p. 41).

Although there have been limited methodical studies conducted to examine the usefulness of data-driven decision making about value-added models beyond Tennessee (TVAAS), in Pennsylvania (PVAAS) and Ohio systems (SOAR) the use of value-added models is still new. Counter to the limited systematic studies, one can find volumes written about how awful the value-added models can be. Many of these articles consider limited research with little evidence showing the use of data to inform “decision making” which may or may not lead to improved school performance and may contribute to student learning (Feldman & Tung, 2001). Value-added models are part of an extensive system. The systems use test to drive reforms with some evidence showing that when trained, teachers develop a deeper understanding of how these measures work. Information from standardized tests can then lead to significant changes in a teacher’s instructional practices and potentially student achievement (Hamilton, 2003; Stecher, 2002; Supovitz & Klein, 2003; Symonds, 2003).

In isolation, numbers generated by a value-added model cannot determine if instructional practices are suitable or detrimental. Determining positives or negatives depends on several factors, including educator capacity to interpret and use data. (Choppin, 2002; Marsh, Pane, & Hamilton, 2006; Saunders, 2006). With several factors influencing outcomes, there is a need to extend professional development beyond instructional content. Adding to the complexity, many end-of-year tests in several value-added model systems make it difficult to use data for instructional purposes. When value-added models are supplemented with benchmark/interim tests administered over the course of the school year immediate feedback is provided to teachers, making data more meaningful (Supovitz & Klein, 2003; Symonds, 2003).

Researchers Lewis and Ruhil (2006) developed an important study using the School's Online Achievement Results (SOAR) assessment in an evaluation of the effects of value-added

model on student outcomes. Using data from 63 Ohio school districts participating in the SOAR projects and several samples of matched comparison districts, Lewis and Ruhil (2006) estimated the effect of receiving the value-added model reports on student achievement outcomes. The study surveyed several similar approaches and used a variety of analytic methods which all led to the same results: There was no evidence that receiving the value-added model reports affected fourth or sixth-grade student achievement in reading, mathematics, science, or citizenship. However, the concern with the Lewis and Ruhil study is that thorough information is not provided on how the district personnel, principals, or teachers used the value-added model data to inform or analyze instructional practices. There is no way to determine the level of integration of the data in staff's decision-making process. The researchers do report that 14 of the 63 SOAR districts were fully implementing the value-added model through "use at the building level throughout the district by principals and teachers to inform decisions about instruction" (Lewis & Ruhil, 2006, p. 5). Interestingly, when Lewis and Ruhil paralleled these districts to matched comparison districts, they discovered the scores for sixth-grade students in these districts trended significantly statistically higher than those in the matched comparison districts.

Interestingly, from the value-added models themselves, one cannot determine the impact without further study of the processes used for instruction and data analysis. Highlighting part of the appeal of value-added models is the expectancy that the process can delineate the effects of schools or teachers from the contextual background individualities of students and thereby afford a stronger measure of effectiveness in the pedagogy and practice of educators in a school or district. This leads to the conversation of whether value-added models provide means of the causal effects schools and teachers have on students (Braun, 2005; Kupermintz, 2003; McCaffrey et al., 2004; Raudenbush, 2004; Reckase, 2004). Still, the question of using data to

determine professional development needs of educators is not at the forefront of these conversations.

While direct demonstration that value-added models are causal would be very challenging, validating that value-added models have causal effects is more straightforward, because the systems may be considered an intervention that can be provided to some schools or districts and not others (Rubin, Stuart, & Zanutto, 2004). While it is evident the establishment of a value-added model system can have positive effects on student outcomes. Research has not determined exactly how this is clear. In many studies, the measurement of all factors in many Value-added model studies has not done. In isolation a value-added model would not decisively show approximations of causal effects of schools or teachers, but it would imply that any faults in the measures are adequately insignificant in the framework of the value-added model. Continually looking at the system in basic terms of a data set of student achievement or a simple construct of a teacher's ability to *move* a student academically will always result in a shallow analysis of what is occurring in a school. Pushing deeper into the construct of a teacher's understanding of the value-added model, professional development centered on data usage, and pedagogy will develop a more complete picture of the viability of a system based on a value-added model.

### **Summary**

"The government is extremely fond of amassing great quantities of statistics. Statistics raised to the nth degree, the cube roots extracted, and the results arranged into elaborate and impressive displays. What must be kept ever in mind, however, is that in every case, the figures are first put down by a village watchman, and he puts down anything he damn well pleases" (Sir Josiah Charles Stamp, 1880–1941).

So what is the real intent of value-added models? Is it only numbers put down by the "village watchman" or does it have a more succinct purpose? Additionally, the government and its entities understand what is being measured, assessed, and potential outcomes of the data. However, the average educator, who is being judged by these systems has little to no understanding of how value-added models work. The systems and student assessment, in general, have become increasingly controversial issues. If you want to measure change, you should not change the measure represents one of the issues with TVAAS (Beaton & Zwick, 1990).

There have been several shifts in the test to further confuse the average educator. However, there is data on both sides of the argument on the effectiveness of methods to determine student growth and teacher effectiveness in the process. In considering all the factors involved in using value-added models to determine teacher success, what level of knowledge the teacher has of the value-added model does not seem to have been taken into account. The lack of teacher understanding as an independent measure indicates a misunderstanding of how to use TVAAS tool as a tool for professional growth. On February 4, 2014, the Tennessee Education Association effectively challenged the use of TVAAS as the exclusive tool for determining teacher effectiveness, pointing out that the score is only an estimate and subject to wide variations due to standard error (Rusk, 2014).

## **Chapter Three:**

### **Methods and Procedures**

During this research, the question explored was: “Why do certain schools have high TVAAS value-added scores?” The purpose of this study was to analyze influences an educator's understanding and knowledge of value-added models, specifically the Tennessee Value-Added Assessment System (TVAAS), has on the value added assessment scores for the schools in which educators work. The researcher examined teacher perceptions and knowledge of TVAAS by looking deeper into the educator’s opinions of the subject matter. This phenomenological study utilized five elementary schools in a local East Tennessee school district focusing on educators in grades three through five with at least one year of experience with TVAAS before the 2014 school year. Educators new to the school district in the 2014 academic year forward have not had consistent experience with TVAAS due to a variety of test formats used and because validity of data from this time frame questionable. The researcher selected this methodology to explore the reasons why some schools have high TVAAS value-added scores and to determine whether an educator’s knowledge and perceptions had an effect on scores. This study utilized qualitative measures using the phenomenological method.

### **Definitions**

TVAAS- Tennessee Value-Added Assessment System- the value-added model used to determine student growth and achievement scores.

Achievement- is the measure of what a student is expected to achieve on TVAAS.

Value-Added- how much a student grew as measured by a value-added model about expected achievement score. This is a specific term referring to the growth measure used in Tennessee

Phenomenological- a category of qualitative research where the researcher explores the lived experiences of individuals during a particular phenomenon (Creswell, 2007; Marshall & Rossman, 2011)

Tennessee Comprehensive Assessment Program (TCAP)- the comprehensive exam used to develop TVAAS scores.

Value-added model- Refers to all generic forms of value-added systems used to measure student growth

### **Research Design**

The amount of research about teacher knowledge and perceptions of TVAAS related to the level of value-added scores at the schools in which they teach is limited. Considering the lack of research, the goal of this study was to explore the relationship between educator's perceptions and knowledge at schools with high TVAAS value-added scores. The study was qualitative in nature.

Qualitative research is a form of scientific research. Scientific research involves an investigation that pursues answers to a question using a systematically predefined set of procedures to respond to the research question. Qualitative research does not forecast what is to happen in the near term; rather, it is an analysis that shows a depth of understanding a concept explored at particular setting and time (Ary, et al., 2013). The process involves the collection of evidence that produces findings not determined in advance and yields results that are applicable beyond the proximate limitations of the study. Qualitative research has these qualities. Also, qualitative research attempts to comprehend a particular research problem from the viewpoints of

the local populace it includes. Qualitative research is particularly efficient in acquiring accurate information about the values, opinions, behaviors, and social contexts of particular populations (Nkwi, Nyamongo, & Ryan, 2001).

The strongest quality of qualitative research is the capacity to provide complex textual descriptions of the experiences of people in a given research issue. It provides knowledge about the humanistic views of a research topic. Qualitative methods are also useful in identifying intangible factors, such as social norms, socioeconomic status, gender roles, ethnicity, and religion, whose role in the research issue may not be readily apparent (Denzin & Lincoln, 2000). When coupled with quantitative methods, qualitative research can interpret and enhance the understanding of the complicated existence of a given research topic. In one way, qualitative research differs slightly from scientific research in general. Findings from qualitative data can often be extended to people with characteristics similar to those in the study population, gaining a rich and complex understanding of a particular social context or phenomenon typically takes precedence over eliciting data generalized to other geographical areas or populations (Bernard, 1995).

Phenomenological studies start with the conjecture that participants' multiple realities are embedded in a participant's perceptions. Thus, an experience has diverse meanings for each. The investigator, through unstructured interviews, researched the participant's thoughts and feelings to obtain the essence of an individual's experience (Ary, et al., 2013). Phenomenology is a science whose purpose is to describe particular phenomena, or the appearance of things, as lived experiences (Streubert & Carpenter, 1999). However, phenomenological research methodology is hard to explain because it has no clearly defined steps. The lack of clearly defined steps is due, in part, to the reluctance of phenomenologists to place primary emphasis on time or sequence of

events. Phenomenologists are of the opinion that the precise definition of methodology tends to limit researchers' creativity (Burns & Grove, 1998).

Phenomenological research is a textual description used to illuminate what the participants experienced. Phenomenology is a proper philosophic method that was utilized to describe the phenomena of elementary educators in schools with high TVAAS scores. Themes were divided into two groups individual and collective. Individual themes were considered in helping to determine what individual experiences shaped responses and may lead to further analysis to see if providing the individual experiences on a larger scale may be beneficial to other schools not included in the study. Collective experiences were used to determine positive or negative trends with TVAAS. Participant responses were utilized to provide the evidence needed to understand the lived experiences of educators in high performing TVAAS schools.

Some of the disadvantages of phenomenological research are: One interviewing requires much more time than questionnaires (Burns & Grove, 1998). Second, because of the sample size time is usually limited (De Vos, 1998). Subject bias is always a threat to the validity/trustworthiness of the findings and inconsistency in data collection from one subject to another (Holloway & Wheeler, 1996). The vast amount of data collected makes ordering and interpretation difficult. Also, unstructured interviews are conducted by the researchers themselves and not by other interviewers (Polit & Hungler, 1995).

While there are some disadvantages, phenomenology offers several advantages beyond the general advantages of qualitative research. First, phenomenology is a highly appropriate approach to studying human experience (Wimpenny & Gass, 2000). Next, it tries to uncover concealed meaning in the phenomenon embedded in the words of the narrative (Sorrell &

Redmond, 1995). Lastly, as a research method, phenomenology is a rigorous, critical, systematic investigation of phenomena (Streubert & Carpenter, 1999).

Within any research area, various contributors can have diverse thoughts. Qualitative samples must be sufficient to guarantee that most perceptions that would be important are discovered. Also, at the same time, if the data set becomes too large, information can become repetitive. If a researcher remains true to the values of qualitative research, sample size in the majority of qualitative studies should generally follow the concept of saturation (Glaser & Strauss, 1967).

Saturation regulates the bulk of qualitative sample size. There are other factors that can determine how fast saturation is achieved in a qualitative research. Charmaz, (2006) advises that the aims of the inquiry are the fundamental driver of the study design and sample. For example, a small study with limited goals focused on a single concept would accomplish saturation quicker than a larger study that looks across multiple concepts. In education, a small study may look at specific groups thought on standardized testing in a specific region, while a larger study may take in multiple groups across varied geographic regions.

Other scholars have also clarified additional influences that can determine a qualitative sample size and saturation in qualitative studies. There are seven factors that may affect the potential size of a sample of a study. One, the heterogeneity of the population. Two, the number of selection criteria. Three, the extent to which 'nesting' of criteria is needed. Four, groups of special interest that require intensive study. Five, multiple samples within one study. Six, types of data collection methods use. Lastly, number seven, the budget and resources available (Ritchie, 2003).

Due to the many factors that can dictate sample sizes in qualitative studies, many researchers are reluctant to suggest what establishes an appropriate sample size. Some researchers find the lack of specifics frustrating. Guest, Bunce and Johnson (2006) say, "Although the idea of saturation is helpful at the conceptual level, it provides little practical guidance for estimating sample sizes for robust research prior to data collection." When conducting a literature review for this study the researcher found only seven first person research papers that provided guidelines for actual sample sizes. Creswell (1998), states for phenomenology research 5 to 25, while Morse (1994) recommends no less than 6 with no upper limit given. Bertaux (1981) recommends 15 as the smallest sample for any qualitative study.

A fourth factor that is likely to influence sample size is the heterogeneity of the population from which the sample is drawn. For some research questions, the population may be quite heterogeneous with a good deal of sub-group variability. It is possible, if not likely, that a researcher will want to capture at least some of this variability in view of the likelihood that it will be associated with significant variability in experiences and world views of participants (Ary, et al., 2013).

### **Researcher**

As an educator with a variety of life experiences: military, private sector sales as well as educational experiences; special education teacher, special education director, and principal (grades k-12), the two constant themes that have always shown results are training and review of data. Training has taken two forms. One has been on-the-job training and the second is related to professional development. Also, data has always been available for all of the experiences mentioned. Whether in the Army with training proficiency data or sales with tracking monthly progress, it is apparent life is full of data.

Upon arrival in Tennessee ten years ago, the Tennessee Value-Added Assessment System (TVAAS) was a foreign concept. It became evident quickly that gaining knowledge of the measure of assessment by which schools and educators were assessed was necessary. It was rapidly evident that TVAAS has two reporting categories and it was important to know how they connect. The first is Achievement, which is the measure of what a student is expected to achieve on TVAAS. The second is value-added or how much a student advanced toward an expected Achievement score. When first taking over an East Tennessee school ten years ago, it had all A's in the four content areas (math, reading, science and social studies) in the achievement category on the state report card. However, the school was in the bottom 10-15 % of value-added scores when compared to the 49 other elementary schools in the school district. Something needed to change and to be efficient in making the change, researching the TVAAS system was critical.

Three aspects of TVAAS quickly emerged. One, teachers had an extreme dislike and mistrust of TVAAS. Many had the opinion that TVAAS reduced their performance to a number that judged them either positively or negatively, thus, completely factoring out the human element of education (mainly student demographics). Two, because of the negative attitude about TVAAS, teachers taught whatever they deemed age appropriate with no connection to the state standards measured by TVAAS. Finally, staff had no concept of best practices. Not one teacher in grades 3-5 (TVAAS tested grades) did small groups or had any real knowledge of current instructional practices.

As the issues of why the school was not performing better (lack of curriculum development, lack of adherence to curriculum and limited understanding of best practices) on TVAAS value-added data became clearer, it became apparent a school-wide improvement plan was needed. A three-year professional development cycle was the cornerstone of this program.

The first phase started with a three-month review of the school district's current curriculum framework. Next, the school moved on to developing an understanding of Bloom's Taxonomy and Marzano's frameworks as connected to best educational practices for elementary school teachers. (Bloom, et al. 1956; Marzano, 2008)

The initial process took two years to implement fully. Lastly, in year three, using the Professional Learning Communities (DuFour, 1999) model and through grade level and vertical planning, staff worked on developing curriculum maps and pacing guides based key indicators tested by TVAAS, as determined by faculty and administration.

In addition to professional development, a theme was developed with staff to get them to further the understand that TVAAS was a measure of how students preformed in relation to the standards. It was simply a tool to help determine what instructional practices did or didn't work. The comprehensive school improvement plan resulted in the school achieving all A's in both achievement and value-added for seven years in a row. For reference, the next closest school in the district only accomplished the same scores for two years in a row. Additionally, the school is currently in the top 50-75 schools in the state depending on the academic subject. Also, the majority of the work by staff on developing curriculum was adopted by the school district. These achievements were accomplished using professional development that addressed the needs of the school, as highlighted by TVAAS data, and working closely with staff to understand their concerns while working to develop a shift in mindset from negative to positive about TVAAS.

Through this research, the goal was to determine whether or not other schools that have achieved high levels of success have experiences similar in nature. If so, how did the educators figure out the steps needed to be successful with TVAAS? If the experiences were not the same, then what process or steps were used to achieve successful TVAAS scores?

### **Study Participants and Setting**

The accessible population for this study will be five public elementary schools in A large East Tennessee school district, with high TVAAS scores for 2012, 2013 and 2014 school years. From 2015 to the current years, the testing methods used have varied and test results do not have the same validity as previous time frames. The participants will include grade three through five teachers and administrators from five schools with the highest value added scores in the school district.

Two sampling methods were used for this study. First, Random Purposeful Sampling, which is frequently used if a potential purposeful sample is too large, in order to enhance the credibility of the study by randomly selecting participants or sites from the larger group (Ary, et al., 2013). This method was used to choose the teachers from each of the four schools selected based on TVAAS scores. The second method used to choose the five schools was Extreme Sampling. This form of sampling selects units that are atypical, special, or unusual. In this case, five schools with high TVAAS value-added scores, at least all A's in value-added for one year, compared to other schools in the county. This type of study can be used to identify practices, teaching methods, and student characteristics that may be relevant to superior performance (Ary, et al., 2013).

### **Data Collection**

The TVAAS value-added data used to select the schools for this study already exists within the data gathered from the Tennessee Department of Education and the school districts Office of Accountability. The school TVAAS data collected was retrieved using the website for each organization and using the test results of 2011, 2012, and 2013 years of TVAAS

assessments in both reading language arts and mathematics. After the study is performed, teacher knowledge and perceptions will be significant to the findings of the research.

The study used phenomenological inquiry through in-depth interviews to obtain the lived experiences of the educators included in the research. The phenomenological approach was used to understand the subjective aspects of educator's frames of reference that achieved high TVAAS scores were the results of individual factors experienced by the educators. "Researchers in the phenomenological mode attempt to understand the meaning of events and interactions of people in particular situations" (Bogdan & Biklen, 2003, p. 23). Personal experiences of the individuals involved in the study were recorded using interviews. Interview questions were designed to draw out meaningful descriptions and clear meaning from participants as they described what they felt led to their school's success with TVAAS. A list of the highest performing TVAAS schools was used as a guide when determining eligible participants for the study. The five schools selected included varying demographics.

The purpose of the research was described and the procedures were clarified for each participant. Confidentiality and anonymity were minimal, as no personal information was disclosed in the research findings and all notes with identifying characteristics were omitted from the record of information. The interview process consisted of one meeting centered on the educator's past experiences with the phenomena, present experience with the events, and perceptions of how they felt.

In the phenomenological method the researcher uses observations, field notes, and recordings as secondary sources of data (Marshall & Rossman, 2011). These sources of data can be used as complementary to the primary sources and used in conjunction with the interviews to

enhancement the primary data. Enhancing this data can assist with data analyzation and is a tactic to show the data interpretations was credible and reliable (Creswell, 2007).

The primary foundation for this study consisted of participant interviews. The researcher used a semi-structured approach. A semi-structured interview is a qualitative method of inquiry that combines a pre-determined set of open questions with the chance for the researcher to explore particular themes further as the discussion develops (Ary, et al., 2013).

### **Analytical Methods**

Phenomenology is a category of qualitative research in which the researcher explores the lived experiences of individuals during a particular phenomenon (Creswell, 2007; Marshall & Rossman, 2011). While there are varied philosophical approaches to phenomenological research, the primary focus of phenomenology research is to describe individual experiences of phenomena. As defined by Van Manen (2002), the term essence does not describe the “whatness” of a phenomenon, but rather the meaning relations that we maintain with the world (Van Manen, 2002). Essence is a relational term that refers to ways of encountering and relating to the things of our world before and while we comprehend them in language and poetic and conceptual thought (Ary, et al., 2013). For this study, an analysis was conducted using horizontalization, or identifying significant statements or quotes, from the in-person interview. Meaning and themes or clusters of information were determined from the respondent’s answers (Ary, et al., 2013).

Out of the fundamental concepts of the various thoughts of phenomenological research involves the extent of researcher subjectivity (Finely, 2012). Generally, for researchers conducting phenomenological research, the researcher should seek to keep an open mind

throughout the investigative process in an attempt to see the experiences in a fresh, new way and to aside personal views and experiences during the research process (Finlay, 2008).

Ary, et al. (2013), outlined four processes of the phenomenological research method. One is bracketing or epoche, the process of the research setting aside bias. The researcher should reflect on personal biases relating about the study in an attempt to bracket, or set aside, preconceived experiences in order to gain a fresh understanding of the experiences of the participants in the study (Creswell, 2007; Marshall & Rossman, 2011; Moustakas, 1994).

The second process is the reduction. This is a phenomenological device that strives to bring characteristics of meaning into nearness or focus. This is not so much a procedure as a certain thoughtful attentiveness and synthesis of the analyzed data (Ary, et al., 2013). In this process, it is important to thoroughly examine all data to remove anything not directly associated to the essence of the experience. As a result, all data is clustered by theme to provide a synopsis of the experience.

Third, through an essential or invariant structure, a description is developed. A textual description is used to explain the experience whereas the structural description clarifies the framework that influenced the occurrence, how it was experienced, and in what conditions and situations. After examining these descriptions and the researcher's experiences, a composite description is written that conveys the overall essence of the phenomenon (Ary, et al., 2013).

The final step in the phenomenological research process involves synthesizing the evidence to gain the essence of the experience (Moustakas, 1994). The researcher writes about the mutual experiences of the contributors to provide the reader with an understanding of what it is like to experience the phenomena (Creswell, 2007; Moustakas, 1994).

## **Ethical Considerations**

At times the trustworthiness of qualitative research is questioned because concepts of validity and reliability, as used in quantitative studies, cannot be expressed in the same way when utilizing qualitative studies. To promote the concepts of reliability and validity in qualitative research, Guba developed four criteria (Lincoln & Guba, 1985). The first is credibility in preference to internal validity. For this study, credibility is established through the choice of using phenomenological research that is a proven method well established both in qualitative research in general and specifically in information science. The next criterion is transferability in preference to external validity. Transferability refers to the degree to which the results of qualitative research can be generalized or transferred to other contexts or settings (Trochim, 2006). While this is approached from the perspective of the person trying to generalize the study, the researcher can enhance the transferability by doing a thorough job of describing the research context and the assumptions that were central to the research (Ary, et al., 2013).

The third aspect is dependability in linking to reliability. The idea of dependability, on the other hand, emphasizes the need for the researcher to account for the ever-changing context within which research occurs. The researcher is responsible for describing the changes that take place in the setting and how these changes affected the way the researcher approached the study (Lincoln & Guba, 1985). The final benchmark is confirmability taking the place of objectivity in qualitative research. Qualitative research tends to assume that each researcher brings a unique perspective to the study (Ary, et al., 2013). Miles and Huberman (1994) consider that an essential criterion for confirmability is the extent to which the researcher admits his or her predispositions. The researcher documented the procedures in this study checking and rechecking the data throughout the study to ensure confirmability.

Researchers should be familiar with ethical issues involved in their research and have the moral obligation to consider the rights of the participants of the study strictly. The three key concepts are consent, protection from harm, and the right to privacy (Ary, et al., 2013). It is important for the researcher to establish trust between the participants and the researcher. Also, it is the responsibility of the researcher to be impartial and protect the anonymity of the participants involved. Additionally, the researcher should put safeguards in place to ensure the participant's rights, needs, and values are at the forefront of the research process ensuring that the participants rights are not harmed during the process (Streubert & Carpenter 1999). Procedures put in place to ensure anonymity include numbers being assigned to participants instead of using names. Written documentation was provided to each participant before the interviews were conducted and each participant signed a document acknowledging voluntary participation.

### **Summary**

The goal of this study was to explore the question: "Why do certain schools have high TVAAS value-added scores?" TVAAS data was used to select schools with high TVAAS value-added scores. Individual participants chosen from each school were interviewed to determine any themes within the school that have led to the high level of success. The researcher examined teacher perceptions and knowledge of TVAAS by looking deeper into each participant's opinions of the subject matter. This phenomenological study utilized five elementary schools in an East Tennessee school district in East Tennessee. Only educators in grades three through five with at least one year of experience with TVAAS before the 2014 school year interviewed. The study used phenomenological research to determine the findings of the research question.

## **CHAPTER 4**

### **Findings**

This study explored the following question: “Why do certain schools have high TVAAS value-added scores?” The intent was to analyze the influences an educator's understanding and knowledge of value-added models, specifically the Tennessee Value-Added Assessment System (TVAAS), has on the value-added scores for a school. Data collection entailed a series of interviews with both administrators and teachers in grades three through five. A total of nineteen interviews were conducted with four principals, one academic coach, and 14 teachers. The academic coach was questioned in place of one principal because the data was gathered from the schools where the researcher is currently the principal. For the purpose of this study, the academic coach is included in the administrator/principal group. The assistant principal for that school conducted the interviews. Participant’s identities were kept anonymous for reporting of the information.

The research focused on five schools in an East Tennessee School Districts that have achieved A’s in value-added in Reading and Math at least once between the academic years of 2012-2015. These schools consisted of two suburban schools with low poverty and minority populations; two city schools with 100% free-reduced lunch populations, minority student rates between 47-86%, and over 50% of the minority population as English as A Second Language (ESL); and one rural school with a 6% free-reduced lunch rate.

The purpose of the phenomenological research method is to expound a deeper understanding of the research question. Scholars have clarified seven factors that can determine a

qualitative sample size and saturation in qualitative studies. The first factor is the heterogeneity or make up of the studies' population. The second factor is the number of selection criteria used when determining participants. A third consideration is the extent to which 'nesting' of criteria is needed. The fourth factor includes the groups of special interest that require intensive study beyond the initial scope of inquiry. The fifth factor is ensuring that there are multiple samples within the study. The sixth consideration is the types of data collection methods used in the research. The seventh and last factor includes the budget and resources available to the research for the study (Ritchie, 2003).

Two sampling methods were used for this study. First was Random Purposeful, which is a sampling method frequently used if a potential purposeful sample is too large. It also is used to enhance the credibility of the study by randomly selecting participants or sites from the larger group (Ary, et al., 2013). This method was used to choose the teachers from each of the five schools selected based on TVAAS scores. The second method used to determine the five schools was Extreme Sampling. Extreme Sampling selects units that are atypical, special, or unusual. In this case, the units are five schools that have high TVAAS value-added scores compared to other schools in the district. This type of study can be used to identify practices, teaching methods, and student characteristics that may be relevant to superior performance (Ary, et al., 2013).

The study was limited to teachers that had been employed in the school district since approximately 2012. Nesting was not a high consideration and special interest groups were not needed to attain accurate data. The data collection method consisted of two sets of interviews. The first was with educators from the five different schools. The second was with the respondents from the school that had the highest value-added scores. Additionally, under the multiple resources category the budget was not a concern as resources were available at no cost.

The sample size and variation led to a robust degree of saturation. The heterogeneity was respectable with both principals and teachers in the various experience levels sampled which also included a wide variety of school demographics.

As defined by Ary, Sorenson, and Walker (2013), in *Introduction to Research in Education*, the process of reduction was used for analysis of the interview data. When using this method, the researcher writes descriptions of the participants' experiences and how those experiences were perceived. Themes were developed from reading, rereading and reading the text again. From the analysis, the researcher derived an overall description of the general meaning of the experience through a process called reduction, sometimes referred to as data condensing. According to Ary, et al. (2013), reduction is a phenomenological device whose goal is to bring facets of meaning into nearness or focus. It is not so much a procedure as a certain thoughtful attentiveness. In analyzing data, the researcher looked at three key points. One, knowledge of what TVAAS measures. Key indicator words would center on student growth. Two, specific types and levels of professional development. Responses would indicate school, district or state level professional development with further indicators being specific themes or names like Marzano (2008), Bloom's Taxonomy (1956) or data analysis. Three, words that were based on emotion to determine any cultural impacts of TVAAS were noted phrases like academic freedom, challenges, and trust. Notes were read, reread and then data coded based on the keywords.

Finally, Member Check was used to verify the accounts of members from one school that had distinct responses to most questions. When using Member Check, the researcher asks the participants in the study to review the findings and conclusions to provide a check on whether the researcher "got the story right." In a peer debriefing, the researcher provides the raw data

along with interpretations to a second investigator (peer) to check whether others would make the same interpretations given the data (Ary, et al., 2013)

### **Demographics**

**School Demographics.** Student population ranged from as few as 400 students to almost 900 (Table 4.1). Two schools had a relatively low number of Free and Reduced students. The number of students on Free and Reduced Lunch program is a general indicator of the poverty level of the school. Two schools had 100 percent of their students enrolled in the Free and Reduced Lunch program. These schools had relatively low student-teacher ratios when compared to the other schools. The most successful school had the highest student-teacher ratio.

Table 4.1

*School Demographics*

School	Enrollment	Minority %	S/T ratio	Free/Reduced %
A	886	25	18:1	15
B	364	8	17:1	6
C	526	47	13:1	100
D	722	19	19:1	23
E	400	86	13:1	100

Administrative demographics for the study were as follows compiled from five administrative interviews. Four principals from elementary schools with grades kindergarten to fifth grade were interviewed. The fifth administrator was the academic coach from the school at which the researcher is the principal. The average experience level was 20.6 years, with the newest administrator having 15 years of experience and the most senior having 26 years of experience as a principal. The academic coach at the researcher's school as interviewed in place

of the principal. This individual has 15 years' experience and has been a member of the administrative team for this project.

Fourteen teachers, from the same schools as the administrators were also interviewed. The average experience for teachers was 9.7 years, with the least amount being three years and the highest being 30. The most common grade level taught was fifth, including eight teachers who currently teach fifth, four teachers are currently teaching in fourth grade, and two teachers are currently teaching in third grade.

### **Data Analysis**

During the data gathering process, a series of ten questions was asked. There was no noticeable difference in responses between those with three years of experience or thirty years. However, there were certain distinctions between the levels of detail in the answers between the administrators and teachers on several questions.

During the research, several themes emerged. Overall, all educators do know that TVAAS is about student growth. High and low achieving students can have positive TVAAS scores in value-added, though they felt gains are harder with high achieving students. Most administrators have attended TVAAS-specific training; but many teachers had not participated in TVAAS professional development. None of the administrators could remember the name of any TVAAS-related events attended. Teacher training generally took place at the school level. There was a mix of district and state training but it was predominately only available to district level administration.

Two administrators and several teachers felt that comparing TVAAS Achievement scores and value-added was not fair and that only value-added should be used as a measure of student growth. While this may be a valid discussion, educators must understand that a school can have

low achievement scores based on the TCAP and have high TVAAS scores. Also, it is important not forget TVAAS Achievement compares a student to a standard. For example, a student who was growing in the category of value-added can be well behind the proficiency levels expected by the state.

Overall, answers provided by the educators involved were not detailed, offered generalities, and did not thoroughly explore concepts. Although not directly stated, principals had, expectedly, a school-wide view of data and student performance. Principals would talk about success with grade levels while teachers mentioned students more often than grade level or school data. The principals had more knowledge and training available while teachers were at the mercy of what was provided by school administrators with some support, on occasion, by the district. It is important to note that the lack of participation by teachers in professional development leads to a knowledge gap not seemingly filled by many principals (Marzano & Waters, 2009).

### **Question One**

The first question inquired: “In your opinion, what was the Tennessee Value-Added Assessment System (TVAAS) designed to measure?” For reference, the state defines TVAAS as follows: The Tennessee Value-Added Assessment System (TVAAS) measures the impact schools, and teachers have on their students' academic progress. TVAAS measures student growth, not whether the student is proficient on the state assessment (TN Dept Ed, 1992).

The key word for this question was student growth. For the first question, 100% of the administrators responded in some form that TVAAS was a measure of student growth. Two principals determined that teacher effectiveness in growing students is a direct outcome of TVAAS data. One principal made a point of highlighting concerns by saying the more you have

to develop complicated formulas to try and improve correlation the less reliable TVAAS can become. When asked for specifics of those formulas, the administrator could not come up with any specific portions of the mathematical formula other than saying it was “complicated.”

Teacher responses were more varied than administration. Nine teachers said that TVAAS was a measure of student growth. Three replied that TVAAS was a measure of the teacher effectiveness. One teacher discussed how TVAAS was connected to teacher observation to come up with an overall effect score. Another response made a more concrete connection saying that “it is also used to see how much academic impact a teacher has had on a student’s learning.” Lastly, a teacher commented that TVAAS measured “the level at which a teacher incorporates effective research-based methods of teaching into a lesson.”

The overall theme of these responses was TVAAS is a measure of student growth. All of the administrators mentioned student growth in their answers. While eleven (80%) of the teachers reported student growth specifically in at least a portion of their answers, three (20%) did not. The three teachers that did not directly mention student growth were indirectly hinting at student growth.

## **Question Two**

The second question asked: “Do you feel students with low-achievement scores can have high value-added scores?” All of the administrators responded that students with low achievement scores could absolutely have high value-added scores. All administrators also felt these students were the easiest to make gains. Notably, two principals, only one of which was from an economically disadvantage school, felt that it was unfair to use low achievement scores against a school “because all that should matter is if the students are growing.” Two principals pointed out that students needed competent teachers to make real growth. Lastly, one principal

remarked that “not only did a student need a skilled teacher but targeted instruction to meet the student’s needs.” All administrators pointed out growth in value-added takes high-level teachers who implement curricula targeted to areas of need as determined by TVAAS data.

All but three teachers echoed many of the same responses as administration. Ten of the respondents felt that did feel students with low achievement could still have high value-added. One teacher answered “probably” but then went on to state “students [that] have low achievement scores can still show growth.” Details in another teacher's feedback demonstrated a depth of understanding about students having to meet or exceed their projected score, saying that students with the “lowest achievement often had the highest level of growth.” This teacher further elaborated that “every student can grow and learn, but there is a break-down when the assessment changes from year to year.” The teacher discussed how one year allowing the read-aloud accommodation on the assessment and then the following year not allowing the read-aloud accommodation. One teacher pointed out the teacher did have an impact but indicated it has to do with connections to the student by saying: “It depends on the situation and the teacher. Some students respond to other teachers. A student might have low achievement scores because of their educational environment.” One teacher felt students needed to be supported both at school and at home before growth can occur with low achievement students. Another cautioned that “some teachers only teach to the test and that leaves gaps in skills at later grades.” Lastly, one teacher said she tends to see the most growth with her students in the lower-mid achievement categories. Overall, the teachers believed this is because the instructional strategies put in place best meet this level of student’s needs. Also noted as important was an individual’s teaching ability, especially when coupled with encouragement toward motivation to learn.

The theme taken from this question is that students with low achievement scores can make high value-added scores. Administrators were more likely to highlight the need for quality instruction and use of data to inform instruction. Several participants from schools with high free-reduced lunch populations felt achievement was a poor measure of these schools and more focus should be on student growth. Even though teachers expressed an understanding of the data, many still reverted to the emotions of teaching by saying that these measures were unfair or results were as much part of teacher “connection” to students. Principals in schools with lower TVAAS scores stated that they had to focus on the individual test, i.e., reading, math, etc., to highlight successes if overall school scores were below expectations. Some would also look at subgroups to see if there were gains in those areas to celebrate the success of these groups. Teachers and administrators felt motivation was important to student growth. Administrators cited teacher motivation while teachers cited student motivation.

### **Question Three**

Question three was the opposite of question two: “Do you feel students with high achievement scores can have high value-added scores?” Administrators’ responses suggested a consensus in this question: “Yes, a little harder from personal experience. If there are weak teachers in prior grades, it helps with having the best teachers in fifth grade who make significant gains.” These same principals worry that their teachers feel these gains are too hard to make. One administrator supporting this concern stated: “Teachers often do not like having these students because they fear they cannot make gains.” Another said: “They often think kids are more ready than they are because of high achievement scores.” All administrators felt that the same process used to grow low students can be used with high achieving students. “Looking at data is a

powerful tool, no matter if the student is low or high achieving.” Wholly, administrators agreed students with high achievement scores can make progress.

The teachers also perceived that students with high achievement can show gains on value added, but they all agreed it is harder to show gains with these students. Most comments were similar to the following: “I think it’s harder to have high value-added for very high-achieving students because they do not have as many gaps to fill (as measured by the former TCAP tests.)” As another noted: “Yes, but it is more difficult because if they are scoring the 95th percentile, it is very hard for the student to grow anymore when they are that close to 100%.”

When reviewing both questions two and three, the key phrases were high gains, high value added, low achievement, and student growth. All of the responses pointed to educators knowing students at both high and lower levels of achievement were able to show value-added growth. All respondents agreed students with lower achievement scores showed gains easier than a student with high achievement scores. Principals tend to look at these two questions from the building and district level, whereas teachers tend to have more of a student and classroom view of student learning.

#### **Question Four**

For question number four, participants were asked if they agreed with the following statement: “TVAAS/TCAP was designed to measure student growth, not whether the student is proficient on the state assessment.” Administrative responses range from a simple yes to more detailed responses that discussed the differences between proficiency and growth, as well as discussing teacher concerns with how the state and district may use student growth data. All administrative responses agreed TVAAS was initially designed to measure student growth, not whether a student was proficient by a standard set by the state department of education. While all

agreed with the above statement, a few principals wanted to point out TVAAS is no longer purely a measure of student performance. As one stated, “I don’t think the intent was to measure teachers in the beginning, TVAAS itself was not meant to measure teachers, but that has morphed into this measure.” Additionally, one administrator wanted to highlight the concern for teachers: “Using data this way has caused a lot of concern for teachers.”

Teachers were not as unanimous in their agreements on this statement as administrators were. While many did give a simple yes, five did not agree 100% with this statement. Three of the five did say they believe it started this way but has morphed into one in the same. One stated: “I do believe that it was designed to do that (measure student growth), but I don’t believe that it is effective in doing so.” Another commented: “No, I do not agree with this statement, they have become one in the same at this point.” Lastly, one teacher said while she did believe that it was designed to measure growth she didn’t “believe that it is effective in doing so because of the all the variables in student’s backgrounds and abilities.”

Two comments stand out and summarize the concerns with TVAAS as viewed through this question. One was that of the principal who wanted to make sure the interviewer was aware of the stress TVAAS may place on teachers. The second was a statement from teacher who said that TVAAS was a measure of student growth that has now changed into a measure of teacher proficiency. This leads the researcher to question why there has been this shift.

### **Question Five**

Question five asked participants if they agreed with the following statement and why or why not: “TVAAS helps educators identify best practices and implement programs that best meet the needs of their students, as well as make informed decisions about where to focus resources while ensuring growth opportunities for all students.”

Administrator responses centered on the use of data to determine strengths and weaknesses of programs. Comments like “Yes, TVAAS helps administrators identify strengths/deficiencies in their school so they can target resources to address student needs” were typical of principals. Specifically, one administrator commented it was the teacher’s “responsibility to make these determinations.” Additionally, one administrator said the following: “The principal provides the guidance and professional development based on analyzing TVAAS data. However, the teachers need to be able to do the work for the results to have more meaning.” This individual clarified, “if the principal does all the analysis and just hands over the results the teachers haven’t been involved in the process and may not trust the results if they don’t understand how people determined the needs of the school, staff and students.”

Like administrators, teachers are also able to utilize TVAAS reports to assist in determining which groups of students were able to make the most progress. This allows teachers to assess the effectiveness of not only strategies, but resources used with each group of students. However, there isn’t overall agreement. Several teachers answered yes, saying that they believe that TVAAS can be used to determine if success an educator and to identify best practices. A key statement was, “the amount of growth that students show lets the educator know if they met the needs of the students.” One staff member replied: “At our school, we do the analysis with our administration and team.” Another teacher indicated that teachers first look at the results on their own then have a grade level data meeting see if there are any common themes in the student growth data. Another reported the administration provided data sheets to use in looking at TVAAS data. The teachers were required to fill out these sheets as a tool to assist in analyzing student scores.

Some teachers, however, were not in agreement. One individual who disagreed with the question commented as follows: “No, since it is a formative assessment with results released in the summer or fall of the following year, it only provides general information. For example, I can look at quintile growth and see that my lowest quintile grew more than my highest quintile. I could then know for the following year that I need to do a better job of meeting the needs of my top students.” The response cited was the better of the disagreement statements with solid reasoning why. Others were more general in their responses but still made the case that TVAAS does not help determine effective strategies: “It does not offer specifics about which instructional strategies are effective or not. The data will tell me I am doing is effective or not...[but] I just don’t know what that strategy is.” Lastly, one teacher who disagreed said, “No one has ever been able to adequately explain the formula to TVASS, therefore, I do not put my trust in something that deals with numbers that no one has ever been able to explain to me.”

The theme emerging from this question is whether as a principal or teacher it is important for the educator to know the limitations of TVAAS. It is only a data point and relies on the end user of the data to determine the reason for the results. The results of this question suggested that the principals have a better grasp of this insight than the teachers. Educators also need to understand TVAAS is limited to showing growth. It is the educators’ responsibility to determine why a student did or did not grow.

### **Question Six**

“What type of feedback do you get about TVAAS scores at the school and district level?” Principals responded that several types of feedback are available. Answers ranged from district level growth data provided by the state to local data like ACT score predictor, scatter plot graphs based on student performance by the teachers, and drill-down reports available for the school and

teacher to help determine best practices. One administrator said “good stuff” is obtained from the TVAAS data pages. Principals also felt teachers get a good amount of data for individuals and school scores. One principal said: “There does seem to be a big gap between special education students and regular education students with lots of feedback for non-special education students.” Although this suggests special education teachers do not always have enough students tested to receive value-added scores, this statement doesn’t equate in the reality of TVAAS reports. Feedback is available at the student level so a special education teacher’s effectiveness could be determined. Also, special education teachers would have access to the same level of data as regular education teachers so special education teachers could do the same data analysis as regular education teachers. Lastly, principals reported there is lots of useful website data available that is used to inform instruction. One principal did caution, “It is a good piece of data, but we make it the whole pie.” One administrator highlighted the importance of making sure staff is receiving appropriate professional development to both interpret TVAAS results and to be able to use the data to make instructional decisions about curriculum.

Teachers, on the other hand, were not confident about the feedback received. Teachers stated that Level 1 teachers get lots of feedback because of poor performance and are the focus of the administration. However, Level 5 teachers do not hear much feedback because they do well and principals don’t spend much time with them on instruction. Another teacher mentioned a similar thought: “Depending on the score it determines the type feedback I need to provide.” The general negative tone continued in another comment: “Almost none, schools write smart goals, and we have a beginning of the year presentations at district and school level, but (what next) questions are not usually approached and attacked throughout a school year.” The last reply was the most negative but within the same tone as many of the other answers. As another teacher

noted, “I get an email saying that my scores are ready and then the school gets a slap on the hand if they aren’t up to par.”

The discrepancy between teachers and principals on this questions is stark. Principals have access to a lot of data and appear to use it in a limited fashion. However, teachers overwhelmingly said they did not get much feedback. While the administrator in one school will talk about how much information is available, teachers at the same school feel that little is available or helpful. Understanding the reasons for this disconnect is important to the process of developing schools that are more effective at growing students.

### **Question Seven**

“Have you had any specific training on how to interpret TVAAS/TCAP scores? If so, was the training provided by the school, district, or state?”

Only two administrators reported attending state level training on TVAAS. The training was several years ago, and neither administrator remembered any specifics. All principals reported regularly attending district level in-service and having access to school level professional development provided by the district to school staff at the principal’s request. All administrators said the informational sessions were generally helpful on how to interpret data. Principals reported the district data days were the “best/most” useful for learning how to use data to improve their schools. It was reported that district data days occur at the beginning of each school year, usually one-two weeks before students return to school. One administrator discussed how the administrative team “interweaved data” into all professional development at the school level. While not all was tied directly to TVAAS data, data was essential in much of what the school did including staff meetings, professional learning communities, and foundational work with curriculum.

Several teachers only responded “no” to this question. Those who did respond overwhelmingly said at best they had a small amount of training by the administration at the school level. One teacher affirmed the value of district data days after having been invited to attend by the administration and felt it was valuable training. Two other teachers referenced specific training coordinated by the principal at their school. The initial training covered how to access and interpret state reports. Curriculum specific training then occurred over the course of three years. One teacher noted that she “received specific training on how to access and interpreter state level reports in addition to specific professional development in the curriculum (district materials, Bloom’s and Marzano) directly linked to TVAAS results.” One teacher replied “yes” said the training “never makes sense.” Another who expanded on the “no” answer responded as follows: “I’ve not had formal training on interpreting TVAAS/TCAP, but it is interpreted for us by our school administration.” Others had equally negative comments. One stated that she had “not seen/been made aware of TVASS training to interpret scores on any level.” Another stated that she had “heard from leaders with comments like, ‘no one knows the weird formula they use,’ and really, none of it makes sense.”

For this question, there was a wide discrepancy of responses between teachers and administration. While administrators had specific professional development offered by the district, it was only provided to teachers on an invitational basis as coordinated by their principal. Also, two administrators had attended training given by the state. Teachers had no training other than what building-level administrators offered. There was a wide variance in training provided at the building level, and the range of responses was interesting. Some teachers received no professional development while others were attending training that covered interpreting results, using the state drill down reports, and specific training in Bloom’s Taxonomy and Marzano’s

instructional strategies to address specific areas of concern. The researcher has wonder what can be done to add continuity of data usage and professional development across multiple schools.

### **Question Eight**

“What influence do you feel an educator’s understanding of TVAAS has on a teacher’s or school’s value added/assessment scores and individual student growth?” Principals expressed the importance of teachers understanding the data to get better results. One stated that “it is very important for a teacher to understand how to interpret TVAAS and the premise behind it.”

Another noted that “if someone can understand the data it allows for deeper reflection and better ownership of the data.” Two indicated they typically see this translate to stronger growth which also translates to more substantial achievement with students. However, cautionary quotes show the added stress teachers can feel from TVAAS data. For example: “Teachers have it in the back of their mind that it (TVAAS Data) is always there,” and “they feel limited to no control and scores are always on teacher’s minds.” A principal did say culture of the school can become negative if scores received aren’t what the school wanted or expected. One administrator did highlight the importance of making sure teachers understand how to use the reports received from the department of education so they could “drill down” through the data to determine strengths and weaknesses of instructional practices from the previous year.

Teachers expressed ideas that showed they understood the importance of TVAAS data. One teacher commented that “having knowledge of our data helps us know how our demographics performed. I believe that it makes us focus more on certain subgroups.” However, even with understanding the importance of data, teachers tended towards emotional responses: “I feel it has little value because the teachers who are called to this profession are going to teach

their hearts out regardless of a score or a number, we tend to look beyond that and go a little deeper than just data.”

Two teachers felt their understanding of TVAAS led to better results because they could use the data in a summative manner to reflect on their instructional practices during the past year. The reflection on data was especially significant when compared to peer’s data in a professional learning community meeting held at the beginning of each year. These two teachers reported their school has them fill out a Beginning of the Year (BOY) goal based on their TVAAS and observation data. This goal is reviewed at the end of the year when the new TVAAS data is available. One teacher outlined a connection between formative to summative data: “There needs to be a combined process that looks how all data is connected.” She further described the process they use. “Currently we work to connect the data we get from benchmark testing to predict how students may do on TVAAS. We had a really good system when we used Discovery Ed to predict TVAAS. The last few years haven’t been good because of a change in benchmark test”.

The opinions of the principals do not align with the thoughts of the teachers for these questions. While principals have many opportunities to attend professional development at the district level these same opportunities do not seem to exist for teachers. Also, principals do not appear to be passing their knowledge on to teachers. However, some teachers from one school did show a better grasp of TVAAS.

### **Question Nine**

“Do you see TVAAS having the ability to influence school culture? If so, in what ways?” Three principal answers focused on the scores a school or teacher receives. Statements such as “some schools have much to celebrate and this makes for a positive culture” and “efforts need to focus on the positive scores to build up the moral” were common among all three. One principal

pointed out if scores are not good staff can feel “deflated.” Administrators further stated if scores are negative you cannot let it dominate your culture. One administrator had a different take on the question: “Yes, depending on the time of year it can be positive or negative. During test taking time it is stressful. However, it is usually positive when we get the results back for most teachers.” In one school that struggles with achievement, the principal makes it a point to highlight the difference between achievement and value-added. Reminding teachers to focus on student growth, the principal also tells staff that scores “are not always a fair reflection of the teacher's abilities, it is a snapshot of one test on one day.” The response that may sum up this question best is from one principal: “Yes, it does, when student growth scores are high and recognized, teachers, are encouraged, motivated, and have self-efficacy. When student growth scores are low and this data is published, it demoralizes the staff and creates a negative culture in the school.”

Teachers shared similar thoughts to principals, i.e., that TVAAS scores can have both negative and positive impacts on school culture. One teacher said, “I think that TVAAS can influence the school culture by being a teacher motivator especially in schools with low achievement scores.” Meaning if a school has strong value-added it can make up for the possible discouragement of low scores in other categories. Particularly insightful, however, was one teacher highlighting how using TVAAS as part of the teacher evaluation can be demoralizing. This teacher focused on how the state changed the norms for generating TVAAS a couple of years ago and how these numbers get linked to teachers for bonuses for high TVAAS scores. The teacher felt the bonus system was unfair. This individual mentioned in some cases, teachers did not get individual growth scores and had to use school ratings. The linking of these results to bonuses led to some teachers who got to use school scores getting a monetary bonus while those

who had individual scores did not. These teachers were in the same grade level. Again, the linking of TVAAS scores to money and evaluations led to poor morale and a lot of dissatisfaction.

Overall, teachers in most schools viewed TVAAS as a negative. Eleven teachers said they saw TVAAS as having a negative impact on school culture with the negative consequences becoming more evident when teachers started seeing the linking of evaluation data with a growth score reducing their performance to a number. Three felt somewhat indifferent about TVAAS and said they pay little attention to it, focusing mostly on formative assessment data and curriculum pacing guides from the school district.

Two teachers, however, said the overall impact at their school was positive, and the negative feeling was generated from the state and district. These teachers believed their principals have made TVAAS about data analysis to improve practice through a system of professional development. The professional development was implemented over the course of several years. The focus on how to use the data to improve instruction, create a competitive environment towards other schools, and use a systematic curriculum improvement approach encouraged academic freedom within the guidelines set by the administration.

### **Question Ten**

“Any other thoughts not covered?”

Only two principals responded to this question. One principal wondered if we are “stretching all kids, both high and low.” This individual said it seems like “a lot of effort is focused on the struggling students and the high students sometimes get left out of the conversation.” Also, this principal felt there were too many changes from the district and the state to stay with one set of ideas for improvement stating: “In the last three years we have

switched district benchmark test, TVAAS/TCAP has been renormed, and the school district has adopted a new reading and math curriculum. Too many shifts to keep staff energized and focused.” The other principal commented, "I do not think that teachers/students should be subjected to a state test that does not look at each student individually and holistically. There are too many factors that determine a student’s learning in one academic year, plus there is a lot of change right now with curriculums within the district.”

Five teachers provided additional comments. One felt too much emphasis was connected to money/jobs causing teachers to take “short cuts” to get the students to achieve better scores by not addressing the needs of the whole child and just teaching to the test. This person did not elaborate on what type of short cuts a teacher could take to get better scores. Another teacher said she did not have an issue with measuring performance but did not like attaching it to evaluations and bonuses. A third echoed similar thoughts: “I like TVAAS for general information. “I do not think they should tie it to evaluation scores...especially when teachers have no direct impact/contact with the students, for example, a 2nd-grade teacher taking the growth score of 5th-grade teacher.”

The fourth teacher wanted to expand on the concept of how professional development can be focused on TVAAS data and the administration using data as a tool has been the biggest factor in not being afraid of TVAAS. She also commented that the link between TEAM and TVAAS scores is uncomfortable, but that is a district and state level issue and is not an issue at the school. The fifth teacher was more direct than the others: “TVAAS is a joke.” Overall, this teacher felt it was unfair to use this measure to “judge” a teacher’s performance. She stated there are way too many issues with the population of students they work with to use a test to determine a teacher’s effectiveness.

## Member Check

Member Check was used at the school selected for the researchers' second round of interviews and conducted with the school whose participants gave the more in-depth answers. This school also had the best TVAAS scores. There were three questions asked. One, what specific professional development activities have you engaged in that has led to the continued success with value-added scores? Two, are there any structures in place that facilitate data analysis? Three, why do you feel the school has had continued success with TVAAS?

The themes reinforced from question one were professional development connected to data analysis about formative and summative data, using Marzano's instructional strategies and Bloom's Taxonomy for best instructional practices, and understanding curriculum and pacing guides with focus on the connection between curriculum and data analysis. The administrator from this school pointed out that "through a professional development process, it became apparent the link between test data and the flow of curriculum has to be related to the other, or there is no foundation for understanding what you did instructional." Another comment that stood out was, "I realized a strong connection existed between formative and summative data (Discovery Ed and Benchmark testing/TCAPS) and being able to use this data to grow students." Lastly, "the use of Marzano's helped take abstract curriculum standards and put some meaning to them. Blooms Taxonomy was important in looking at higher order teaching about Marzano's and curriculum standards."

There were two themes for question two. One, the process was methodical and took place in small steps over the course of several years. One teacher said, "It seems like the data analysis cycle never stops!" The administrator discussed how in the beginning, data analysis was very rudimentary: "In the beginning we started with simple math module test at grades 3-5 for

analysis, it was the baby step needed to make stronger connections to what happened in the classroom as measured by a test.” The point was to give a particular test that teachers frequently use as the starting point of how to look at data and then stretch this concept to TVAAS data. Lastly, the administrator stated, “I think the results were an effect of a systematic approach to curriculum implementation, an active mentoring system by veteran staff for new teachers, and a well-developed PLC system that focuses on formative assessments result to inform instruction.”

For question number three the overall theme was competitiveness, both internally and in comparing the school’s performance to other schools. One staff member cited the high performance of all staff: “The staff here is so high powered that you have to push yourself to keep up with the group, there is no one on top in all categories, and best performer will often shift from year to year.” Another teacher pointed to the principal showing the staff other schools’ performance. “The principal was always challenging us to have better scores than the schools close to us. The data presentations would often include the name of the schools and the principal always said we could do better than them, and we DID!” Also mentioned was the principal’s “hold harmless” approach. “The principal has worked hard to make us feel at ease when looking at data, good or bad; I always felt it was a tool to help me become a better teacher.”

### **Summary**

For the research process, a set of ten questions was developed to guide the interview process. The questions were constructed in a way not to lead respondents to specific answers but to allow them to focus on several concepts such as student growth, professional development, or school culture. A follow-up interview was conducted with one specific school based on its responses to the questions and the school’s success with TVAAS.

Five schools were included in the research. Two were suburban schools with low Free and Reduced Lunch population, two were city schools with high Free and Reduced Lunch populations, and the last was a rural school with a 6% Free and Reduced Lunch. A total of 19 interviews with four principals, one academic coach, and 14 teachers occurred.

All participants had at least three years of experience and worked in their current school during the TVAAS data years surveyed. Each had the basic conceptual understanding that TVAAS measures student growth. Administrators typically provided lengthier answers. The length of answers was a factor of more training on how to understand and use TVAAS data. Other than one school, no other group could define any program or sequential effort to determine TVAAS data usage or connections to the curriculum.

The researcher purposefully avoided leading questions like “What type of professional development have you done at your school about TVAAS?” While direct questioning may have made the data collection process easier, it could have led to biased answers by leading the participants to respond in ways tailored to the wording of a particular question.

The answers of participants from one school resulted in a follow-up interview with the three educators from this school. Noted by teachers from this school were themes such as academic freedom, professional development, goal setting, and data-driven instruction within the supports of a curriculum framework. Also, educators from this school highlighted the collegial environment where data builds pedagogy and was not a potential negative influence over a teacher’s career.

While specific professional development for TVAAS, beyond understanding how to read TVAAS reports, was not discussed regularly in isolation, these educators linked several other school level initiatives to the school’s success. This school continually achieved high value-

added scores. The follow-up interviews resulted in a distinct pattern of behavior at this school not noted in the other schools involved in the study.

Teachers also discussed goal setting. An important part of the of the process was setting goals based on data: “Our SMART Goal (Specific, Measurable, Agreed Upon, Realistic, Time Sensitive) process for the grade level was important to make us look at data across the grade level and set a common goal.” Another teacher appreciated the chance to set a goal based on individual performance: “Sitting down with the administration and reviewing my performance in a collective manner helps me feel more focused in what I need to accomplish during the upcoming year.” In the researcher’s experience developing SMART goals and requiring teachers to set individual goals was not unique to the highest performing school in the study. However, this school started the process of goal setting at least two years earlier than the other schools in the study.

## **Chapter Five**

### **Conclusions, Inferences and Recommendations**

During this study the research question explored was: “Why do certain schools have high TVAAS value-added scores?” This study analyzed the influences of an educator's understanding and knowledge of value-added models, specifically the Tennessee Value-Added Assessment System (TVAAS), has on the value-added scores for a school.

Three schools achieved all A’s on the TVAAS value-added categories of reading and Math for at least one year during the period survey. One school managed to do it two years in a row in all four subjects surveyed by TVAAS: Reading, Math, Science, and Social Studies. The fifth school has achieved all A’s in all subjects for seven years in a row. This school is the only school that could refer to specific processes that led to its success.

Student assessment is both the beginning and end of the educational process. In truth, whether an evaluation is meant to be formative or summative, the data derived must be used to advance instructional abilities, thus growing students. Research has suggested that less effective schools have more variation in the performance of teachers when compared to high performing schools. Assessment is needed to determine if a student is growing and to ascertain the best instructional strategies. “Teacher candidates must learn not only content, pedagogy, learners, and communities. They must know themselves as educators. Also, given the diversity in schools, teacher candidates must acquire and sustain individual beliefs and reflect those same thoughts through demonstrated professional dispositions.” (Harrison, Smithey, McAfee, & Weiner, 2006).

TVAAS data is an important tool for schools to use in understanding student growth. This was of particular importance during the follow-up interviews. Participants pointed out how it developed into the focal point for all other analysis. Teachers at one school became adept enough using benchmark data (Discovery Ed testing) that they could, in a general sense, determine how well a student would do on TCAPs. Data at this school was used to shift the pedagogy of teachers and to make the school more successful in growing students.

This reaffirms the news that has emerged from previous studies that look at ways to improve student achievement. Two different comprehensive syntheses of research on the factors impacting student learning have come to the same conclusion: the most important variable in the achievement of students is the quality of instruction they receive on a daily basis (Marzano, 2003; Hattie, 2009). Be it in a classroom, school or central office, the focus on assessment and student data is clear. The introduction of legislation like No Child Left Behind and other educational measures of student achievement in the emergence of value-added model (think TVAAS) look to push educators to perform at the highest level. This pressure, when combined with the establishment of standards-based education with concepts like Common Core, the pressure is on to show student growth.

### **Conclusion**

The following themes emerged during research. First, educators understand that TVAAS is related to student growth. They also know that students with both high and low achievement scores can have positive TVAAS scores in the value-added category. Most educators believe that value-added scores are harder to attain with students who have high achievement scores. Administrators have attended TVAAS specific training with a mix of district and state training. The state training was several years ago, however, and the administrators who attended could not

remember the name of or specific topics covered. None of the teachers had attended TVAAS-centric professional development at the state level. For all but a few teachers who were selected by their administration to attend a professional development program called data days, professional development for teachers takes place at the school level. At the school level, TVAAS instruction was provided in staff meetings and occasionally during Professional Learning Community meetings. Most discussions of TVAAS were limited to school performance.

The only question that had a disparity revealed in answers between high Free-Reduced Lunch program schools and low was question number two. “Do you feel students with low achievement scores can have high value-added scores?” Two administrators and several teachers responded that comparing achievement and value-added was not fair in these schools. Both the teachers and principals in those two buildings felt only value-added should be reviewed because all that matters is student growth. They felt achievement becomes an unfair measure for most students in schools with these types of profiles. While it is accurate that a school can have low achievement scores based on the TCAP but have very high value-added scores, it should not be forgotten achievement is comparing a student to a standard. The standard, in this case, is a proficiency level determined by the state department of education. Even a student who is demonstrating growth may still be well behind the proficiency levels expected by the state. Thus, proficiency levels should not be overlooked as a measure of student performance.

Overall, the answers provided by the teachers were lacking detail. Teacher responses contained generalities and did not provide in depth specifics even when asked. These general responses result from a lack of exposure to more than surface level TVAAS data and professional development provided at the school level. Meeting with a principal from the district

may have caused some discomfort when providing responses. This could be the fear of providing a response that seems uneducated or inappropriate. Administrators provided more detailed responses and were more at ease during the interview process. Although not directly stated, principals had a school-wide view of data and student performance. Principals talked more about the success of the school or grade levels while teachers were more likely to mention student success. Principals had more knowledge and training available while teachers were at the mercy of what was provided by the schools. The higher standard of understanding by the principals led to a knowledge gap between administration and teachers. This gap does not appear to be filled by many principals with school level professional development. Just as a teacher is a leader, or more importantly a facilitator of learning in a classroom, he or she must also be a facilitator of learning for the school.

The importance of being the facilitator of learning became evident in the follow-up interview with one school. This school experienced a considerable degree of success in both the achievement and value-added categories on TVAAS for several years. It was evident this school had developed a systematic process for addressing its needs highlighted by analyzing school data. The process developed and used by this school allowed educators to isolate curriculum development issues through data analysis leading to a direct impact on student growth. School leadership developed these skills by first modeling the process of data analysis and then facilitating educator learning through Professional Learning Communities.

By isolating instructional issues with data, the principal discovered several potential problem areas. The first was that staff had a lack of understanding as how to interpret data. To remedy this, the principal presented a simple data set with a math module test using a pre/post-test model to demonstrate data analysis. At the end of the math unit grade levels held meetings to

answer several fundamental data analysis questions using the module test data. The following questions were used in the analysis. One, did we meet our goals? Two, in what areas did we do well? Three, what are we most proud of about these results? Four, which results were disappointing or not what we expected? Five, what seemed to work well? Lastly, what strategies will we start, stop, or continue to improve our results?

Once staff developed a basic understanding of data analysis, a second professional development session was conducted to explain TVAAS and value-added data analysis. Staff members were then asked to apply their understanding of data analysis to TVAAS data using the same questions they used for the math module test. During the professional development session, and then over the course of several grade level professional learning community meetings, staff members learned to compare the school results to that of the district and other schools in the district.

One of the themes that was evident in the the belief of teachers that it was more difficult to show positive value-added scores with students who have high achievement scores. This was evident in the beginning of this schools struggle to gain better results. Teachers blamed high achievement results of the students on low value-added scores. As in the research, teachers expressed the same concerns that student achievement scores were too high to show value-added growth. However, analysis of surrounding districts revealed several schools in the region had high achievement and high value-added.

However, even after all the work with data analysis, the school still did not achieve the desired results the first year. After a few months of observation, another series of data analysis, and several conversations with teachers, it was evident the majority of the school was not following district curriculum. Many teachers taught content they determined was important and

paid little to no attention to pacing or curriculum guides. For example, one teacher spent an hour a day during a four-weeks period teaching Roman History to 5th graders, even though it was the four weeks prior to TCAPS and Roman history was not in any elementary or middle school curriculum. The principal set a non-negotiable standard of teaching the curriculum and required the academic coach to add pacing to the agenda for grade level meetings. Once this final piece was implemented, value-added scores continued to rise over the next two years.

Staff turnover retirements, maternity leaves, and promotions lead to a need to reevaluate the overall competency level of a school's staff. During such a review, the data, on a teacher-by-teacher basis, revealed evidence that showed the newest teachers were struggling. New staff was consistently in the 2-3 range on TVAAS with veteran staff continuing to maintaining scores of 4 and 5. It was evident the new staff had not been involved in the professional development critical to the veteran's staff success. New staff members were provided with training in Marzano's Instructional theories and Bloom's Taxonomy, with particular focus on instructional strategies. Also, the master schedule was reworked to provide common grade-level planning time at least one day per week. Monthly meetings to monitor progress and integration of new staff was also included. In addition, the instructional coach led professional development on Bloom's Taxonomy and Marzano's Instructional theories for all staff as both new and refresher training.

### **Implications**

Is the lack of TVAAS knowledge at the school level a result of principals making assumptions teachers know more than they do? A deeper question is did teachers receive basic instruction on data analysis? As educational leaders what is the real role of the principal in supporting teacher use of data? Is it simply to be a manager of information and issues with

responsibilities to help teachers collect, monitor, and interpret the data on individual students, maintain discipline and deal with difficult parents?

An educational leader of a school is a facilitator of adult and student learning who works to promote a positive culture and manage the demands in a modern educational environment. Daily activities include analyzing student data, observing classes, managing budgets and implement policy. However, the most important activities are managing change of the organizational structure by setting high curriculum standards and building effective teams that have strong data analysis capabilities. Weick used the concept of enactment (1979, 2009) to indicate that people in organizations understand the environment by playing an active role in constructing their environment. Data illustrates administrators have a deeper knowledge of TVAAS use and they have a responsibility to ensure that knowledge is shared in a meaningful way to strengthen their organizations. Also, it is important to note that the lack of participation by teachers in professional development leads to a knowledge gap that seems unfilled by many principals (Marzano & Waters, 2009).

### **Recommendations**

First, further research to include schools from outside the district that have similar demographics and results to the schools in this research would be useful to see whether the same patterns of responses exist. During this research, there was evidence that a structured program of data analysis and a complementary professional development program stood out at one school, as did the school's results. The school had highest performance in the district for multiple years and was also in the top 100 performing schools in the state. Progress started with a basic look at data at the classroom level and the program continued to develop through a combined system of understanding data and connecting it to curriculum. Thus, another aspect that should be explored

is the effective use of Professional Learning Communities (PLCs) for both analyzing data and building school culture. It is recommended that other schools implement strategies developed by this school.

Professional Learning Communities can be powerful tools to refine a school's focus, act as a change agent for major pedagogical shifts, promote the growth of both staff and students and decrease the gaps in learning between sub-groups of students. Forming teams of educators that build instructional capacity, while providing guidance from administration, trusting teacher's professional judgment to monitor and adjust curriculum, as well as utilizing methods of instruction to improve student learning, are key to successful Professional Learning Communities.

In addition to the discussions focused on data and best practices, goal setting was also an important practice for teachers. Teachers at the highly successful school were asked to set grade level SMART goals (see Appendix B) and individual goals. At the beginning of the year, the principal met with each teacher to set goals for both TVAAS and TEAM observation based on data from the previous year. Teachers were required to fill out a goal-setting sheet (see appendix A) and this goal was reviewed at the end of the year. Goals were the catalyst for teacher selection of professional development opportunities for the year and provided focus for the year. Other schools should develop SMART goals and the corresponding processes for the work centered on the SMART goal process.

The way information is disseminated by the district and state should also be considered. There is a definite gap between the knowledge and positive perception of TVAAS between administration and teachers. This gap can be directly attributed to the lack of exposure teachers have had with TVAAS. All participants understood the purpose of TVAAS, which was to

determine student growth but from levels of understanding and programs in place varied highly. This is the responsibility of the state, district and school leadership. As capacity is built for teachers-leaders who can promote the goals of the school and build on the ethical standards set in place, the process gets accelerated. Thus, it is recommended including school personnel from the beginning in establishing these ideals allows to have a higher level of ownership in the school community.

This study addressed the question, “Why do certain schools have high TVAAS value-added scores?” The study determined that developing a systematic process for analyzing student data in connection with a strong understanding of curriculum based in a well-defined professional development program significantly increased the chance of a school being successful in terms of student growth. All of the schools involved in the study, only one had such a system in place. There are teachers who have the same depth of knowledge as administrators, but these educators were not the norm in the study. All administrators had the background knowledge to start developing a data-rich system to analyze student growth. However, none spoke of looking at data or curricula systematically and the pace of change in education has led to many starts and stops for many educators.

Administrators had fewer concerns relating to using data than teachers. Administrators had a better understanding of TVAAS than teachers. This is a result of the lack of professional development accessed by or provided to the teachers when compared to administrators. Teachers relied on the school level professional development where administrators regularly attended district level training and occasionally state level training. The district and state level trainings appear more organized and focused. Therefore, it may be beneficial for administrators to begin

conducting more professional development on TVAAS and connections to curriculum at the school level for staff.

The research suggested that a systematic process of data analysis connected to curriculum development can lead to shifts in pedagogy. This shift is carried out in PLCs and further developed through targeted professional development, which leads to deeper knowledge of TVAAS. When these components are combined with best practices, curriculum development, and a positive view of the teacher evaluation system, a culture that promotes healthy competition in a collegial data-rich environment will create strong student growth on TVAAS as measured by value-added.

## References

- Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sense making to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, 66, 136-149.
- Anderson, R. H. (1993). Clinical Supervision: Its history and current context. In R. H. Anderson & K. J. Snyder (Eds.), *Clinical supervision: Coaching for higher performance* (pp. 5-35). Lancaster, PA: Technomic.
- Psychology Health Workplace Program (2012) retrieved from <https://www.apa.org/news/press/releases/phwa/workplace-survey.pdf>
- Ary, D., Jacobs, L., Sorensen, C., & Walker, D., (2013-01-01). *Introduction to Research in Education* (Page 285). Cengage Textbook.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman
- Beaton, A. & Zwick, R. & National Assessment of Educational Progress, Princeton, NJ. (1990). *The Effect of changes in the national assessment disentangling the NAEP 1985-86 Reading Anomaly. Revised*. [Washington, D.C.]: Distributed by ERIC Clearinghouse, <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accn=ED322216>
- Bernard H. (1995). *Research methods in anthropology*, Second Edition. London: Sage Publications, 1995.

- Bertaux, D. (1981). From the life-history approach to the transformation of sociological practice. In Daniel Bertaux (Ed.), *Biography and society: The life history approach in the social sciences* (pp.29-45). London: Sage Publications.
- Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York: David McKay Co Inc.
- Briggs, D., Domingue, B. (2011). Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District teachers by the “*Los Angeles Times*.” National Education Policy Center.
- Brookhart, S. M. (1993). Teachers’ grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123-142
- Bloom, M. (2016) Retrieved from <http://www.ajc.com/s/news/school-test-scores/#about>
- Braun, H. (2016). Getting value out of value-added.: National Academies Press, 2010. Pro Quest E-book Central. Web. 16 October 2016.
- Campbell, D. (1976). Assessing the impact of planned social change.: The Public Affairs Center, Dartmouth College, Hanover New Hampshire, USA. December 1976
- Chaplin, D., Gill, B., Thompkins, A., & Miller, H. (2014). Professional practice, student surveys, and value-added: Multiple measures of teacher effectiveness in the Pittsburgh Public Schools (REL 2014–024). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Childers, T. (2013). The flaming hoops of TVAAAS.: Retrieved from:  
<http://www.timchilders.com/2013/11/12/the-flaming-hoop-of-TVAAS/>.
- Choppin, J., (2002). "Data Use in Practice: Examples from the School Level," paper presented at the Annual Conference of the American Educational Research Association, New Orleans, La., April 2002
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1996). Teachers' assessment practices: preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179
- Coburn, C. (2001). Collective sense making about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23, 145-170.
- Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press
- Collins, C., & Amrein-Beardsley, A. (2014). Putting growth and value-added model on the map: A national overview. *Teachers College Record*, 116(1).
- Collins, C. (2014). Houston, we have a problem: Studying the SAS Education Value-Added Assessment System (SAS® EVAAS®) from teachers' perspectives in the Houston Education Policy Analysis Archives Vol. 20 No.12 24 Independent School District. Arizona State University, Tempe
- Corcoran, Sean P (2010) Can teachers be evaluated by their students' tests scores: Should they be: The use of value-added models of teacher effectiveness in policy and practice;

Education Policy for Action Series, Annenberg Institute for School Reform at Brown University (NJ1)

Corcoran, S., & Goldhaber, D. (2013). Value added and its uses: Where you stand depends on where you sit. *Education Finance and Policy*, 8(3), 418-434.

DOI: 10.1162/EDFP\_a\_00104.

Cranton, P. (2006). *Understanding and promoting transformative learning: A guide for educators of adults*. San Francisco: Jossey-Bass

Creswell, J.W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. 2<sup>nd</sup> ed. Thousand Oaks, CA: Sage Publications.

Cruickshank, D. R., & Haefele, D. L. (1990). Research-based indicators: Is the glass half-full or half-empty? *Journal of Personnel Evaluation in Education*, 4(1), 33–39.

Darling-Hammond, L. (2012) Value-added evaluation hurts teaching. *Education Week*, 31, (24), pp. 24,32, March 14, 2012

Denzin N., Lincoln Y. (2004). *Handbook of qualitative research*. London: Sage Publications, 2000

Derringer, P. (2010). RTT in Tennessee: Assessment done right. *Technology and Learning*, 31(1), 40. Retrieved from <http://www.techlearning.com/article/31572>

Deslandes, R., Rivard, M. C., Joyal, F., Trudeau, F., & Laurencelle, L. (2009). Family-school collaboration in the context of learning assessment practices and communication. Paper presented at the annual meeting of the American Educational Research Association, San Diego.

- DuFour, R., Eaker, R. E., & National Educational Service (U.S.). (1999). *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, Ind.: National Educational Service.
- Eckert, J. M., & Dabrowski, J. (2010, May). Should value-added models be used for performance pay? *Phi Delta Kappan*, 91(8), 88-92
- Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, 66(1), 1-26.
- Elmore, R. F., & Fuhrman, S. H. (1995). Opportunity-to-learn standards and the state role in education. *Teachers College Record*, 96, pp. 432-457.
- Feldman, J., and Rosann T. (2001). "Whole School Reform: How Schools Use the Data-Based Inquiry and Decision Making Process," paper presented at the 82nd Annual Meeting of the American Educational Research Association, Seattle, Wash., April 2001.
- Garland, S. (2012, February 7). Tennessee teacher evaluation systems have rough road ahead. *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/2012/02/07/tennesseeteacherevaluat\\_n\\_1260790.html](http://www.huffingtonpost.com/2012/02/07/tennesseeteacherevaluat_n_1260790.html)
- Glaser, B. & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine Publishing Company.
- Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D., & Whitehurst, G. J. (2011). *Passing muster: Evaluating teacher evaluation systems*. Washington, D.C.: Brookings Institution.

- Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Washington D.C.: Brookings Institution.
- Goldhaber, D. & Hansen, M. (2010). Is it just a bad class? Assessing the stability of measured teacher performance. CEDR Working Paper 2010-3. Seattle, WA: University of Washington.
- Goddard, R. D., Hoy W. K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact of student achievement. *American Educational Research Journal*, 37(2), 479-507.
- Guest, G., Bunce, A., & Johnson, L. (2006). "How many interviews are enough? An experiment with data saturation and variability". *Field Methods*, 18(1), 59-82.
- Hamilton, L.S. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25-68.
- Hamilton, L.S., Stecher, B.M., Marsh, J., McCombs, J.S., Robyn, A., Russell, J., Naftel, S., & Barney, H. (2007). *Implementing standards-based accountability under No Child Left Behind: Responses of superintendents, principals, and teachers in three states*. Santa Monica, CA: RAND.
- Hanushek, E. (1972). *Education and race*. Lexington, MA: D.C. Heath and Company.
- Hanushek, Eric A. (1971). "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-Data," *American Economic Review*, 61(2), 1971, pp. 280-288
- Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. San Francisco: John Wiley and Sons.
- Harris, D. (2011). *Value-added models in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.

- Harris, D.N., and Sass, T. (2005). Value-added models and the measurement of teacher quality. Paper presented at the annual conference of the American Education Finance Association, Louisville, KY, March 17-19.
- Harrison, J., McAfee, H., Smithey, G., & Weiner, C. (2006). Assessing candidate disposition for admission into teacher education: Can Just Anyone Teach? *Action in Teacher Education*, 27(4), 72-80.
- Herndon, Terry. (1985). *Merit pay and the concerns of the teaching profession*. p. 93-98. Lanham: University Press of America, Inc., 1985
- Hill, David, (2000), He's got your number. *Teacher Magazine*, May 2000, pp. 42-47
- Holland, P. (2002). Two measures of change in gaps between the CDFs of test score distributions. *Journal of Educational and Behavioral Statistics*, 27 (1), 3-17.
- Houston Independent School District, (2014). [http://portal.battelleforkids.org/Aspire/Value Added/Progress/Achievement](http://portal.battelleforkids.org/Aspire/ValueAdded/Progress/Achievement).
- Hoy, W.K., & Woolfolk, A.E. (1993). Teachers' sense of efficacy and the organizational health of schools. *Elementary School Journal*, 93, 355-372.
- Ingvarson, L., Meiers, M. & Beavis, A., (2003) "Evaluating the quality and impact of professional development programs". Retrieved from: [http://research.acer.edu.au/professional\\_dev/3](http://research.acer.edu.au/professional_dev/3)
- Jacob, B. & Stephen, L. (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics CXVIII* (3): 843-878
- Jennings, N. E. (1996). *Interpreting policy in real classrooms: Case studies of state reform and teacher practice*. New York, NY: Teachers College Press

- Jette, D., Grover, L., & Keck, C. (2003). A qualitative study of clinical decision making in recommending discharge placement from the acute care setting. *Physical Therapy, 83*(3), 224-236.
- Johnson, M., Lipscomb, S., Gill, B., Booker, K., & Bruch, J. (2012). Value-added models for the Pittsburgh Public Schools. Report to Pittsburgh Public Schools, Cambridge, MA: Mathematica Policy Research.
- Ketelaar, E., Beijaard, D., Boshuizen, H. P. A., & Den Brok, P. J. (2012). Teachers' positioning towards an educational innovation in the light of ownership, sense-making, and agency. *Teaching and Teacher Education, 28*, 273-282.
- Kleiman, S. (2004) Phenomenology: to wonder and search for meanings. *Nurse Researcher 11*(4): 7-19
- Koretz, D. (2001), New Work on the Evaluation of High-Stakes Testing Programs. Symposium at the National Council on Measurement in Education (NCME) Annual Meeting, Seattle, WA., 2001.
- LeClaire, B. (2011, June 1). Will EVAAS® make Wake schools better? Raleigh Public Record. Retrieved from <http://www.raleighpublicrecord.org/news/2011/06/01/will-evaas-make-wake-schools-better-part-ii/>
- Lewis, M., and Anirudh, R., (2006), "Achievement gains study of SOAR pilot and match districts," The Voinovich Center for Leadership and Public Affairs, Ohio University, Athens, Ohio, 2006.
- Lincoln, Y. and Guba, E., (1985). *Naturalistic inquiry*. Beverly Hills, Sage Publications, 1985
- Lipsky, M. (1980). *Street level bureaucracy: Dilemmas of the individual in public services*. New York, NY: Russell Sage.

- Lockwood, J., McCaffrey, D., Hamilton, L., Stecher, B., Vi-Nhuan and Martinez, F. (2006). The sensitivity of value-added teacher effect estimates to different mathematics achievement, The RAND Corporation July 6, 2006
- Marzano, R. J., & Kendall, J. S. (2008) *Designing and assessing educational objectives: Applying the new taxonomy*. Thousand Oaks, CA: Corwin Press.
- Marzano, R. J., & Waters, T. (2009). *District leadership that works: Striking the right balance*. Bloomington, IN: Solution Tree Press.
- McCaffrey, D. F., & Hamilton, L. S. (2007). Value-added assessment in practice: Lessons from the Pennsylvania value-added assessment system pilot project. Santa Monica, CA: RAND Education.
- McCaffrey, D.; Lockwood, J. (2008). Missing data in value-added modeling of teacher effects. *Ann. Appl. Stat.* 5 (2011), no. 2A, 773--797. doi:10.1214/10-AOAS405.  
<http://projecteuclid.org/euclid.aos/1310562205>.
- McCaffrey, D., Sass, T., Lockwood, J., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, 4(4), 572–606.
- McMillan, J. H., Workman, D., & Myran, S. (1999). Elementary teachers' classroom assessment and grading practices. Paper presented at the annual meeting of the American Educational Research Association, Montreal
- Meyer, J. W., & Rowan, B. (1978). The structure of educational organizations. In M. Meyer (Ed.), *Environments and Organizations* (pp. 78-109). San Francisco: Jossey-Bass.
- Miles, M. and Huberman, A, (1994). *Qualitative data analysis: an expanded sourcebook*, 2nd ed. California: Sage Publications.

- Mitchell, C., & Sackney, L. (2009). Sustainable improvement: Building learning communities that endure. Rotterdam: Sense.
- Marshall, C. & Rossman, G.B. (2011). *Designing qualitative research*. 5<sup>th</sup> ed. Thousand Oaks, CA: SAGE Publications.
- Morse, J. (2000). Determining sample size. *Qualitative Health Research*, 10(1), 3-5.
- Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage Publications.
- Murnane, R.J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.
- Nkwi P., Nyamongo I., & Ryan G., (2001). *Field research into social issues: Methodological guidelines*. Washington, DC: UNESCO, 2001.
- Overman, Stephenie, (2012). Fighting the stress of teaching to the test, <https://www.nea.org/tools/fighting-stress-teaching-to-Test.html>
- Polloway, E. A., Epstein, M. H., Bursuck, W. D., Roderique, T. W., McConeghy, J. L., & Jayanthi, M. (1994). Classroom grading: A national survey of policies. *Remedial and Special Education*, 15, pp.162–170.
- Price, T. 2014. Teacher education under audit: Value-added models, TVAAS, EdTPA and evidence-based theory *Citizenship, Social and Economics Education*, 13 (3) pp. 211-225.
- Ravitch, D. (2012). *No student left untested*. The New York Review of Books. <http://www.nybooks.com/blogs/nyrblog/2012/feb/21/no-student-left-untested/>
- Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and selecting samples. In Jane Ritchie & Jane Lewis (Eds.), *Qualitative research practice. A guide for social science students and researchers* (pp.77-108) Thousand Oaks, CA: Sage

- Rose, L. C. and Gallup A. M. (2007). Kappa/Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 89(1), 33-48. Retrieved October 14, 2007, Retrieved from <http://www.pdkintl.org/kappan/kpollpdf.htm>;
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics*, 125(1), 175–214.
- Rowan, B. A. (1982). Organizational structure and the institutional environment: The case of public schools. *Administrative Science Quarterly*, 27, 259-279.
- Rusk, M. (2014). Tennessee tried to mess with teachers and teachers fought back – and won. Retrieved from: <http://www.alternet.org/education/tennessee-tried-mess-teachers-and-teachers-fought-back-and-won>: June 18, 2014
- Sanders, W. & Horn, S. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8(3), pp. 299-311
- Sanders, W., Wright, S., Rivers, J., & Leandro, J. (2009). A response to criticisms of SAS® EVAAS®. Retrieved from [http://www.sas.com/resources/asset/Response\\_to\\_Criticisms\\_of\\_SAS\\_EVAAS\\_11-1309](http://www.sas.com/resources/asset/Response_to_Criticisms_of_SAS_EVAAS_11-1309).
- Sanders, W. (2006). House Committee on education and the workforce hearing on “No Child Left Behind: Can growth models ensure improved education for all students” July 27, 2006. Retrieved from: <http://archives.republicans.edlabor.house.gov/archive/hearings/109th/fc/nclb07270sanders.htm>

- Schochet, P., & Chiang, H. (2013). What are error rates for classifying teacher and school performance using value-added models? *Journal of Educational and Behavioral Statistics*, 38(2), 142–171.
- Sergiovanni, T. (1994). *Building communities in schools*. San Francisco: Jossey-Bass.
- Starratt, R. (2008). Supervision of instruction – the history of supervision, roles, and responsibilities of supervisors, issues trends, and controversies. Retrieved from <http://education.stateuniversity.com/pages/2472/Supervision-Instruction.html>
- Educational Improvement Act, Tennessee §49-1-603, (1992).
- Stecher, B. (2007). “Consequences of large-scale, high-stakes testing on school and classroom practice” in Laura. S. Hamilton, Brian. M. Stecher, and Stephen. P. Klein, eds. *Making Sense of Test-Based Accountability in Education*, Santa Monica, Calif.: RAND Corporation, MR-1554-EDU, 2002, pp. 79–100. As of August 15, 2007:
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), 5-14
- Strout, W., (2013). School level factors associated with Pennsylvania high schools that exceed PVAAS growth projections in reading. The Pennsylvania State University, The Graduate School, Retrieved from [https://etda.libraries.psu.edu/files/final\\_submissions/8828](https://etda.libraries.psu.edu/files/final_submissions/8828) 2016
- Supovitz, J., Valerie K. (2007) mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement, Philadelphia, Pa: Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education, November 2003. As of August 15, 2007:
- Tanner, D., & Tanner, L. (1987). *Supervision in education*. New York, NY: MacMillan Publishing.

- TEA (2009). TVAAS has changed the discussion. Retrieved from <http://www.teateachers.org/news/why-tea%E2%80%99s-position-tvaas-has-changed-first-top-discussions-2009-10>
- Timperley, H. (2011). *Realizing the power of professional learning*. New York: Open University Press.
- Trochim, W. (2006). Qualitative validity-social research methods. Retrieved from <http://www.socialresearchmethods.net/kb/qualval.php>.
- Wainer, H. (2011). *Uneducated Guesses: Using Evidence to Uncover Misguided Education Policies*. Princeton University Press. Retrieved from <http://www.jstor.org/stable/j.ctt7rjsm>
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11(1), 57-67.
- Yorks, L., & Marsick, V. (2000). Organizational learning and transformation. In J. Mezirow (Ed.), *Learning as transformation: Critical perspectives on a theory in progress* (pp. 253-281). San Francisco: Jossey-Bass.
- Zimmerman, B.J. (1995). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 21, 3-8.

## Appendices

## Appendix A

### Individual Data Feedback

### TEAM Observation & Selection Criteria

Teacher Name: \_\_\_\_\_

TEAM Evaluation Summative: \_\_\_\_\_

Individual TVAAS: \_\_\_\_\_

Please refer to the chart below to identify your selection for the TEAM observation cycle. Observations will begin the week of September 8, 2014.

Selections for the 15% & 35% as well as the TEAM observation choice must be completed online by September 2, 2014. Beginning of the year conferences will be schedule next week. There will be a sign up list with Kristy starting Friday. If possible, please try to sign during a time that will not require coverage. If a time slot is not available and you will need coverage please indicate if you will need coverage on the sign up form. Meetings should last about 10 minutes.

Attached is a copy of the Beginning of the Year conference form.

#### Section 1: Results/Trends from Quantitative Measures

-In this section you will need to analyze your student data. The quantitative/growth measure is what you based your 35% measure on (TVAAS data).

-You should look for any noticeable trends in the student performance that may help guide teacher growth plans and area of refinement for the year. At this point, most everyone has multiple years of TEAM and TVAAS data. Look for any trends that maybe emerging.

Section II: As a result of this data review you should be able to set a measurable goal for this school year. The goal may be modified at the BOY conference.

Refinement instructional goal should be a combination of TEAM score and quantitative

data. If you cannot find any correlation between TEAM and quantitative measure focus on quantitative data to develop your goal for this year.

First year apprentice teachers or teachers who do not have TVAAS data, you will need to pick a goal based on an area you need to work on with your classroom instruction. In addition, please remember you will also need to make your 15% and 35% selections for this year by September 2.

SMART GOAL:

When your goal is tangible you have a better chance of making it specific and measurable and thus attainable.

Example: I will elicit higher-order thinking skills by asking questions from the creating, evaluating, and analyzing categories of Bloom's Taxonomy on three out of four days of reading groups to have a "Total Reading" percent of 75-80% at/above the national percentile for proficient and advanced on Stanford 10.

The intent behind a goal should be based on SMART goals.

Specific Measurable Attainable Realistic Timely

Specific: A specific goal has a much greater chance of being accomplished than a general goal.

To set a specific goal you must answer the six "W" questions:

\*Who: Who is involved?

\*What: What do I want to accomplish?

\*Where: Identify a location.

\*When: Establish a time frame.

\*Which: Identify requirements and constraints.

\*Why: Specific reasons, purpose or benefits of accomplishing the goal.

EXAMPLE: A general goal would be, “Get in shape.” But a specific goal would say, “Join a health club and workout 3 days a week.”

---

Measurable - Establish concrete criteria for measuring progress toward the attainment of each goal you set.

When you measure your progress, you stay on track, reach your target dates, and experience the exhilaration of achievement that spurs you on to continued effort required to reach your goal.

To determine if your goal is measurable, ask questions such as.....

How much? How many?

How will I know when it is accomplished?

Attainable – When you identify goals that are most important to you, you begin to figure out ways you can make them come true. You develop the attitudes, abilities, skills, and financial capacity to reach them. You begin seeing previously overlooked opportunities to bring yourself closer to the achievement of your goals.

You can attain most any goal you set when you plan your steps wisely and establish a time frame that allows you to carry out those steps. Goals that may have seemed far away and out of reach eventually move closer and become attainable, not because your goals shrink, but because you grow and expand to match them. When you list your goals you build your self-image. You see yourself as worthy of these goals, and develop the traits and personality that allow you to possess them.

---

Realistic- To be realistic, a goal must represent an objective toward which you are both *willing* and *able* to work. A goal can be both high and realistic; you are the only one who can decide just

how high your goal should be. But be sure that every goal represents substantial progress. A high goal is frequently easier to reach than a low one because a low goal exerts low motivational force. Some of the hardest jobs you ever accomplished actually seem easy simply because they were a labor of love.

---

Timely – A goal should be grounded within a time frame. With no time frame tied to it there's no sense of urgency. If you want to lose 10 lbs, when do you want to lose it by? "Someday" won't work. But if you anchor it within a timeframe, "by May 1st", then you've set your unconscious mind into motion to begin working on the goal.

Your goal is probably realistic if you truly *believe* that it can be accomplished. Additional ways to know if your goal is realistic is to determine if you have accomplished anything similar in the past or ask yourself what conditions would have to exist to accomplish this goal.

T can also stand for Tangible – A goal is tangible when you can experience it with one of the senses, that is, taste, touch, smell, sight or hearing.

## Appendix B

### Grade Level SMART Goal

SMART Goal Example: Using text response in writing that require citing text evidence.

We can use the state writing rubrics as the measurement of student performance with our baseline being this past years TVAAS/State writing assessment. Progress will be monitored using quarterly writing prompts and MIST writing site. Overall there will be a 30% change of students moving from level 2 to level 3. We will administer weekly test response questions which require citing text evidence as part of the process to build capacity in writing. The measurement tool will be the State writing rubrics with anchor papers used as our guiding documents for scoring final works.

Appendix C  
Initial Interview Questions

1. In your opinion, what was the Tennessee Value-Added Assessment System (TVAAS) designed to measure?
2. Do you feel students with low-achievement scores can have high value-added scores?
3. Do you feel students with high achievement scores can have high value-added scores?
4. Do you agree TVAAS/TCAP was designed to measure student growth, not whether the student is proficient on the state assessment?
5. Do you agree with the following statement, TVAAS helps educators identify best practices and implement programs that best meet the needs of their students, as well as make informed decisions about where to focus resources to ensure growth opportunities for all students?
6. What type of feedback do you get about TVAAS scores at the school and district level?"
7. Have you had any specific training on how to interpret TVAAS/TCAP scores? If so, was the training provided by the school, district, or state?
8. What influence do you feel an educator's understanding of TVAAS has on a teacher's or school's value-added/assessment scores and individual student growth?
9. Do you see TVAAS having the ability to influence school culture? If so, in what ways?
10. Any other thoughts not covered?

## Appendix D

### Member Check Interview Questions

1. What specific professional development activities have you engaged in that has led to the continued success with value-added scores?
2. Are there any structures in place that facilitate data analysis?
3. Why do you feel the school has had continued success with TVAAS?