

THE RELATIONSHIP BETWEEN GROWTH SCORES AND THE OVERALL OBSERVATION RATINGS FOR TEACHERS IN A PUBLIC SCHOOL SYSTEM IN TENNESSEE

Joshua Davis

Sullivan County (Tennessee) School System
Blountville, Tennessee

James H. Lampley

East Tennessee State University
Educational Leadership and Policy Analysis
Johnson City, Tennessee

Virginia Foley

East Tennessee State University
Educational Leadership and Policy Analysis
Johnson City, Tennessee

ABSTRACT

The purpose of this study was to investigate the relationship between the TVAAS growth score given by the Tennessee Department of Education and the overall Tennessee Educator Assessment Model (TEAM) observation rating for teachers in grades 3 through 8. The participating county public school system for this study is located in Northeast Tennessee. Participants were teachers in the school system teaching Math, English/Language Arts, Science, and Social Studies in grades 3 through 8 in 10 elementary schools, 6 middle schools, and 2 K-8 schools. Specifically, this research examined the relationship between the TEAM observation scores and overall TVAAS growth score given to the teacher from the Tennessee Department of Education based upon yearly-standardized test scores. Research reinforced mixed views about the validity and purpose of teacher evaluation systems and the use of Tennessee Value-Added Assessment System. Five research questions guided this study and quantitative data were analyzed using a Pearson correlation, one-way MANOVAs and a one-way ANOVA. Results indicated a moderate positive relationship between a teacher's TEAM observation score and the TVAAS growth score given by the Tennessee Department of Education.

INTRODUCTION

The Tenth Amendment to the United States Constitution states: "The powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people." (Tenth Amendment, 1791, para. 1). Because education is not mentioned in the Constitution, it is one of those powers reserved to the states. However the Federal Government has increasingly become involved in public education. The Elementary & Secondary Education Act of 1965 (ESEA) was a Great Society program enacted by the U.S. Congress. The ESEA allocated federal funds for primary and secondary school education. This Act also provided a vehicle to hold schools and states accountable for student achievement (Elementary and Secondary Education Act, 1965).

Public Law 107-110, also known as the No Child Left Behind Act of 2001 (NCLB), was a revision and reauthoriza-

tion of the ESEA (Public Law 107-110, 2002). The stated purpose of NCLB was a fair, equal, and significant opportunity to obtain a high-quality education. The NCLB law called for children to obtain proficiency on challenging state achievement measures.

The American Reinvestment and Recovery Act (ARRA) of 2009 was enacted by the U.S. Congress to stimulate the economy, support job creation, and invest in critical sectors including education. The ARRA created a platform for educational reform by calling for adoption of standards and assessments, measurement of growth and success, measures to improve teacher quality, and improvement of low-achieving schools (U.S. Department of Education, 2009). The ARRA provided 4.35 billion dollars in a Federal grant program known as Race to the Top. Tennessee was announced as one of the first states to receive Race to the Top grant funds. Tennessee's application, titled First to the Top, included reforms to cur-

riculum standards, new assessment measures, and a new teacher evaluation system (Tennessee Consortium, 2013).

After passage of the First to the Top legislation, Tennessee adopted the Tennessee Educator Assessment Model (TEAM) for teacher evaluations. The TEAM model was different from previous models because it called for an increase in frequency of observations and indicators for teacher performance (Tennessee Educator, 2014). Teachers are also linked to student performance to determine teacher effectiveness through the Tennessee Value-Added Assessment System (TVAAS) and the Tennessee Comprehensive Assessment Program (TCAP). For the first time, Tennessee teachers were given an effectiveness rating determined by observation scores (TEAM), student effect data derived by state assessments for achievement (TCAP), and academic growth (TVAAS).

STATEMENT OF THE PROBLEM

The Tennessee Department of Education now links TEAM observation scores and student achievement data (TVAAS and TCAP). At the conclusion of the 2012-2013 school year, the Department of Education released data for every public school in the state with a number, 0-5, stating how closely TEAM observation scores related to student achievement and growth data according to the Tennessee Value-Added Assessment System. Additional research is needed to assess the relationship between teacher growth scores and teacher observation scores. The purpose of this study is to determine if there is a relationship between the TVAAS growth score given by the Tennessee Department of Education and the overall TEAM observation rating for teachers given by system administrators in grades 3 through 8 in a Tennessee school system.

RELATED LITERATURE

Teacher Evaluation in Tennessee

Tennessee's new performance-based teacher evaluation model (TEAM) requires administrators to rate a teacher's performance on lesson planning, classroom environment, lesson standards and objectives, student motivation, lesson structure and pacing, teacher questioning, teacher content knowledge, teacher knowledge of students, the grouping and arrangement of students, academic feedback, activities and materials, student thinking, and student problem solving. The evaluation system requires 50% of the evaluation to be comprised of student achievement data that includes 35% based on student growth measures represented by the Tennessee Value-Added Assessment System (TVAAS) and 15% based upon additional student achievement measures selected by the teacher (Tennes-

see Department of Education, 2012). Observation scores through the state's TEAM model comprise the other 50% of the evaluation.

Beginning in the summer of 2011 Tennessee provided training for principals and system administrators who would be evaluating teachers. These administrators were required to pass an inter-rater reliability exam in which administrators viewed a video of lessons being delivered by teachers and rated teachers on the TEAM rubrics to ensure that administrators understood the different rating levels of rubric (Tennessee Department of Education, 2012).

Implementation of the evaluation system began with the start of the 2011-2012 school year. As implementation continued through the first semester of the 2011-2012 school year, it became clear that satisfaction with the evaluation system varied considerably from district to district (Tennessee Department of Education, 2012). As a result of negative public reaction to the teacher evaluation system, Tennessee's Governor assigned a panel, the State Collaborative on Reforming Education (SCORE), with the task of conducting an independent review of the evaluation system, including collecting feedback from every school district across the state (Tennessee Department of Education, 2012). As a result of the SCORE report, there were several changes made in the second year of the TEAM evaluation system in Tennessee. Additionally, there was increased district flexibility through the approval of more than 40 plans to further customize the overall evaluation system to fit the needs of individual districts (Tennessee Department of Education, 2013).

Tennessee Value-Added Assessment System

The Tennessee Value-Added Assessment System (TVAAS) was created in 1992 as a component of the Education Improvement Act (Tennessee Department of Education, 2014). TVAAS is based on SAS's Education Value-Added Assessment System. TVAAS is a statistical method that is designed to measure the impact schools and teachers have on their students' academic progress. The TVAAS method uses previous test data to plot a growth pattern for every student in grades three through 12 in Tennessee. Growth is measured by how much gain or progress an individual student or group of students make over time. Under Tennessee's teacher evaluation legislation, value-added scores count for a portion of teachers' overall evaluation scores (Tennessee Department of Education, 2015). Sanders and Horne (1998) stated that the TVAAS model, along with other measures including promotion, attendance, and dropout rates of individual schools, would provide information to create a new system of accountability for Tennessee schools.

Supporters of the value-added models emphasize potential for improving student achievement. Jerald (2009) stated "value-added data provides principals, teachers, and parents with valuable information about students' past and predicted performance and give teachers feedback about the effectiveness of their own classroom instruction" (p.2). Value-added proponents assert that tracking of student achievement and the use of value-added data can help teachers and administrators to meet the individual needs of students.

Some researchers have argued that the TVAAS model does not control for socioeconomic status (SES) and demographic factors that can affect the starting point in student achievement and the rate at which a student learns (Darling-Hammond, 1997; Kupermintz, 2002; Linn, 2001; University of Florida, 2000a; University of Florida, 2000b). Additional criticism has surfaced indicating concern over the use of value-added data to determine teacher tenure, pay, and decisions relating to the continuation of employment (Berliner, 2013; Konstantopoulos, 2014; Yettick, 2014).

In Tennessee a teacher's growth score is calculated by the Tennessee Department of Education. The TVAAS or growth score indicates the amount of growth students assigned to the teacher have demonstrated on state TCAP tests during that testing cycle. Teachers receive ratings of 1 through 5 based upon the percentage of students demonstrating at least 1 year of growth. A teacher whose students have demonstrated 1 year of growth received a score of 3 indicating that the teacher has met the standard. Scores less than a 3 are considered below the standard and scores greater than a 3 are considered exceeding standards.

SAMPLE AND DATA COLLECTION

Two hundred forty teachers in grades three through eight were selected for inclusion in this study. These teachers were selected because they receive teacher effect data based upon courses they are directly responsible for teaching. The teachers were assessed and observed during the 2012-2013 school year with data being reported in September 2013. The participating school system is comprised of 23 schools and serves over 10,000 students in grades Pre-K through 12.

Teacher observation scores and teacher growth scores are reported to school systems by the Tennessee Department of Education annually. After obtaining approval of the Director of Schools from the participating school system, the data were coded by the school system to protect the anonymity of the teachers prior to releasing information. No identifying information was associated with the growth or observation scores used in this study.

METHOD

This nonexperimental, quantitative study was conducted using a secondary data analysis design. Inferential statistical analyses (Pearson correlation coefficient, ANOVA, MANOVA) were used to answer 5 research questions (Green & Salkind, 2011). IBM-SPSS statistical software was used for the analysis of the data and an alpha level of .05 was used to determine statistical significance.

LIMITATIONS AND DELIMITATIONS

It was assumed that the data that were collected from the state's databases were valid and reliable. It was also assumed that the methodology adequately addressed the research questions. In addition, it was assumed that the statistical tests were appropriate and possessed the necessary power to detect, if present, differences in the variables. This study was delimited to teachers who teach in grades three through eight in the participating school system in Tennessee. Teachers who met all other qualifications but did not had both a TVAAS growth score and a TEAM observation score were excluded from the study. This study was further delimited by the theoretical framework that was selected for the research. The results may not be generalizable to other school systems or other states.

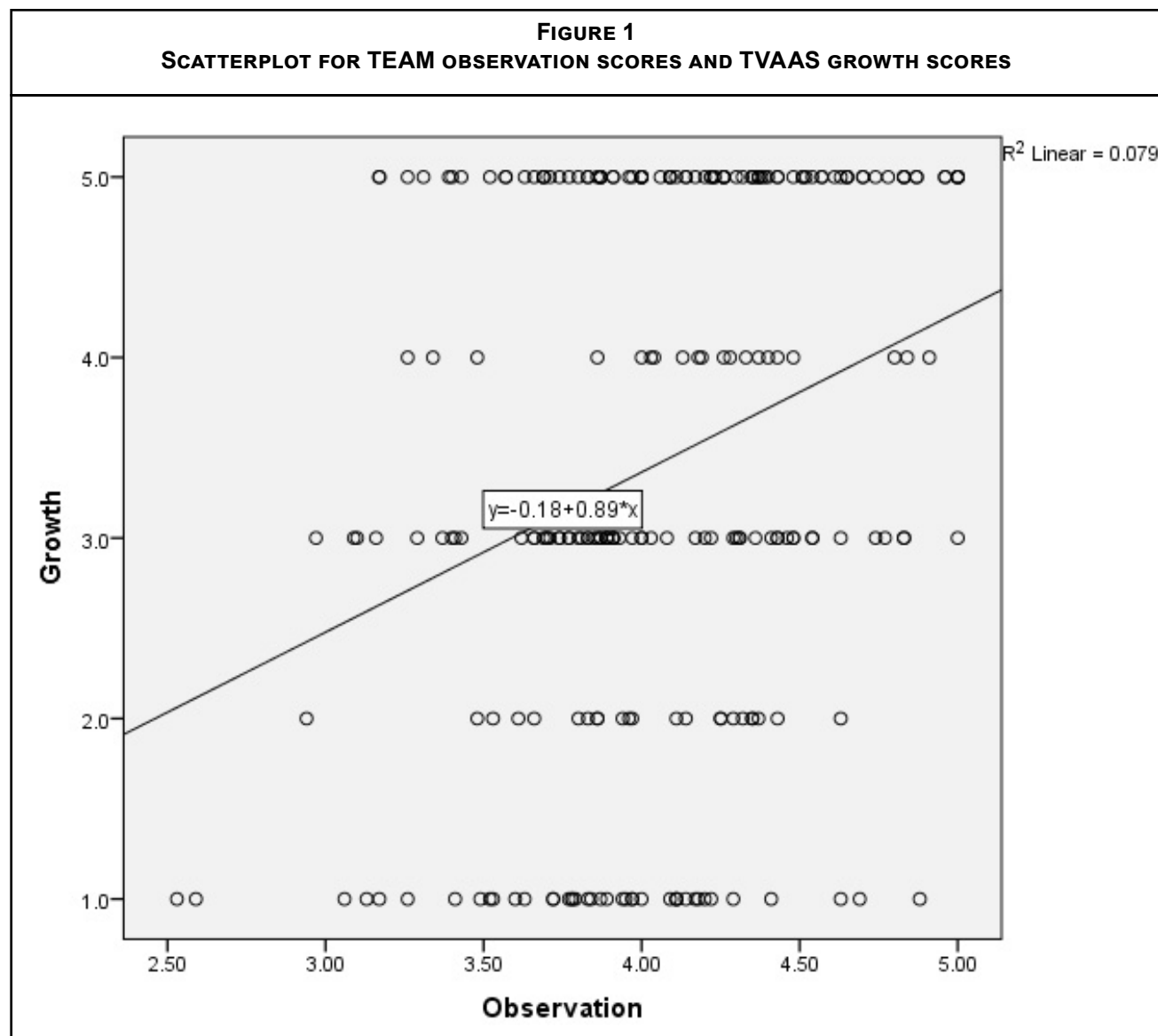
FINDINGS

Research Question 1: Is there a significant relationship between overall TEAM observations scores and TVAAS growth scores given by the Tennessee Department of Education for teachers in grades 3 through 8 in the participating school system?

A Pearson correlation coefficient was computed to test the relationship between TEAM observation scores and TVAAS growth scores. The results of the correlational analysis revealed a moderate positive relationship between TEAM observation scores ($M = 4.05, SD = .47$) and TVAAS growth scores ($M = 3.41, SD = 1.49$) scores and a statistically significant correlation [$r(238) = .28, p < .010$]. In general, the results suggest that teachers with high TVAAS growth scores tended to have high TEAM observation scores. Figure 1 displays the bivariate scatterplot.

Research Question 2: Is there a significant difference in teachers' TVAAS growth scores and teachers' TEAM observation scores by gender for teachers in grades 3 through 8 in the participating school system?

A one-way multivariate analysis of variance (MANOVA) was conducted to determine the relationship of the gen-



der (male or female) of the teacher to the two dependent variables, TVAAS growth scores and TEAM observation scores. There was no significant difference in TVAAS growth scores or TEAM observation scores between male teachers and female teachers, Wilks' $\Lambda = .98$, $F(2, 476) = 2.40$, $p = .090$. The multivariate η^2 based on Wilks' Λ was .02. Male and female teachers tended to have similar TVAAS growth scores and TEAM observation scores. Table 1 contains the means and standard deviations on the dependent variables and the gender of the teacher.

		TEAM Observation		TVAAS Growth	
Gender	N	M	SD	M	SD
Male	42	3.91	.07	3.35	.23
Female	198	4.08	.03	3.42	.10

Research Question 3: Is there a significant difference in teachers' TVAAS growth scores and teachers' TEAM observation scores by license type for teachers in grades 3 through 8 in the participating school system?

A one-way multivariate analysis of variance (MANOVA) was conducted to determine the relationship of license types (apprentice or professional) to the two dependent variables, TVAAS growth scores and TEAM observation scores. A significant difference was found for license type and the dependent variables, Wilks' $\Lambda = .94$, $F(2, 237) = 7.56$, $p = .001$. The multivariate η^2 based on Wilks' Λ was .06. Table 2 contains the means and standard deviations on the dependent variables of license type.

An analysis of variance (ANOVA) was conducted on each of the dependent variables (observation and growth) as follow-up tests to the MANOVA. Using the Bonferroni method, each ANOVA was tested at the .025 level (.05/2). The ANOVA for license type and observation scores was found to be statistically significant, $F(1, 202) = 9.72$, $p = .002$, $\eta^2 = .04$, and the ANOVA for license type and growth scores was also statistically significant, $F(1, 34) = 9.35$, $p = .002$, $\eta^2 < .038$. Teachers holding professional licenses tended to have higher TVAAS growth scores and higher TEAM observation scores than teachers holding apprentice licenses.

		TEAM Observation		TVAAS Growth	
Type of License	N	M	SD	M	SD
Professional	204	4.09	.46	3.53	1.43
Apprentice	36	3.83	.51	2.72	1.67

Research Question 4: Is there a significant difference in teachers' TVAAS growth scores and teachers' TEAM observation scores by socioeconomic status of the school in grades 3 through 8 in the participating school system?

A one-way multivariate analysis of variance (MANOVA) was conducted to determine the relationship of the school's socioeconomic status (Title I or Non-Title I) to the two dependent variables, TVAAS growth scores and TEAM observation scores. There was no significant difference in TVAAS growth scores or TEAM observation scores between Title I schools and Non-Title I schools, Wilks' $\Lambda = .99$, $F(2, 476) = .58$, $p = .557$. The multivariate η^2 based on Wilks' Λ was .01. Teachers in Title I and non-Title I schools tends to have similar TVASS growth scores and TEAM observations scores. Table 3 contains the means and standard deviations on the dependent variables of socioeconomic status of the school.

		TEAM Observation		TVAAS Growth	
Socioeconomic Status	N	M	SD	M	SD
Title I	186	4.06	.51	3.47	1.49
Non-Title I	54	4.05	.31	3.22	1.53

Research Question 5: Is there a significant difference in teachers' TEAM observation scores among the 4 levels of experience of the evaluating administrator (0-1 year experience, 2 to 4 years experience, 5 to 10 years experience, 11 or more years experience)?

A one-way analysis of variance (ANOVA) was conducted to evaluate the relationship between the evaluating administrator's experience and the overall TEAM observation rating. The factor variable, years of experience, included four levels (0-1 year of experience, 2 to 4 years of experience, 5 to 10 years of experience, 11 or more years of experience). The dependent variable was the overall TEAM observation rating. The ANOVA for experience of administrator in observation scores was significant, $F(1, 238) = 11.96$, $p < .001$. The strength of the relationship between the experience of the administrator and the observation rating, as assessed by η^2 , was large (.13).

Because the overall F test was significant, post hoc multiple comparisons were conducted to evaluate pairwise difference among the means of the four groups. A Dunnett's C procedure was selected for the multiple comparisons because equal variances were not assumed ($p = .006$). There was a significant difference in the means between administrators with 11 or more years of experience and all three of the other groups. However, there was not a significant difference between the means of any of the other pairs. It appears that administrators with more experience award higher observation scores. The 95% confidence intervals for the pairwise differences, as well as, the means and standard deviations for the four groups are reported in Table 4.

SUMMARY

Important findings for this study included a moderate positive correlation between teachers' TEAM observation scores and their TVAAS growth scores, no significant difference on TEAM observation scores or TVAAS growth scores by gender of the teacher, a significant differ-

TABLE 4
MEANS AND STANDARD DEVIATIONS WITH
95% CONFIDENCE INTERVALS OF PAIRWISE DIFFERENCES

Administrator's Years of Experience	N	M	SD	2-4 Years	5-10 Years	11 or more Years
0 – 1	24	4.04	.28	-.31 to .08	-.26 to .15	.13 to .61
2 – 4	71	3.93	.38		-.11 to .24	.27 to .71
5 – 10	102	3.99	.51			.20 to .65
11 or more	43	4.14	.46			

ence on TEAM observation scores and TVAAS growth scores between teachers holding a professional license and those holding an apprentice license (teachers with professional licenses scored higher), no significant difference in TEAM observation scores and TVAAS growth scores between Title I and non-Title I schools and, and a significant difference in TEAM observation scores among the levels of experience of the observing administrator (more experienced administrators tended to award higher observation scores).

CONCLUSION

The state of Tennessee has made changes over the past several years in the way schools and teachers are evaluated. These changes are in response to legislation that has led to an increased emphasis on testing and accountability. Changes to Tennessee's teacher evaluation model include the adoption of the Tennessee Educator Accelerator Model (TEAM) and the incorporation of student achievement and growth data (TVAAS) for teachers' overall annual evaluations. Further research is suggested to examine other public school systems in Tennessee to determine if the results are specific to the participating public school system. There was also concern noted regarding the Tennessee Value-Added model and the validity of producing student growth scores from state achievement measures with a formula that is unknown to the public. The data set for this study was important because it was produced prior to the announcement by the Tennessee Department of Education that it would report a correlation score for the teacher TEAM observation scores and student TVAAS growth scores for each school. The Tennessee Department of Education's expectation is that a school's TEAM observation scores should be positively correlated with the school's TVAAS growth scores. A numerical score is now given by the Department of Education to every school noting this correlation. The most important question may be whether or not these state generated correlation scores will produce an artificial alignment of TEAM observa-

tion scores and TVAAS growth scores by influencing future TEAM observation scores.

REFERENCES

- American Reinvestment and Recovery Act (ARRA). (2009). Retrieved from <http://www.gpo.gov/fdsys/pkg/BILLS-111hr1enr/pdf/BILLS-111hr1enr.pdf>
- Berliner, D. C. (2013). Problems with value-added evaluations of teachers? Let me count the ways. *Teacher Evaluation*, 48(4), 235-243.
- Darling-Hammond, L. (1997). Toward what end? The evaluation of student learning for the improvement of teaching. In J. Millman (Ed.), *Grading teachers grading schools. Is student achievement a valid evaluation measure?* (pp.248-263). Thousand Oaks, CA: Corwin.
- Elementary and Secondary Education Act. (1965). Retrieved from <http://www.ed.gov/esea>
- Green, S. B., & Salkind, N. J. ((2011). Using SPSS for windows and macintosh: Analyzing and understanding data. Boston, MA: Prentice Hall.
- Jerald, C. D. (2009). The value of value-added data. *The Education Trust*, 11.09.
- Konstantopoulos, S. (2014). Teacher effects, value added models, and accountability. *Teacher College Record*, 116(1), 1-21.
- Kuppermintz, H. (2002). *Teacher effects as a measure of teacher effectiveness: Construct validity considerations in TVAAS*. CSE Technical Report 563. University of California, Los Angeles.
- Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems*. CSE Technical Report 539. University of California, Los Angeles.
- Public Law 107-110 (No Child Left Behind). (2001). Retrieved from <https://www.k12.wa.us/esea/NCLB.aspx>

Sanders, W. L., & Horne, S. P. (1998). Research findings from the Tennessee value-added assessment system database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247-256.

Tennessee Consortium on Reforming Education. (2013). *Educator evaluation in Tennessee: Initial findings from the 2013 first to the top survey*. Retrieved from <http://www.tnconsortium.org>

Tennessee Department of Education. (2015). *Tennessee Value-Added Assessment System*. Retrieved from <http://www.tn.gov/education/data/TVAAS.shtml>

Tennessee Department of Education. (2012). *Teacher evaluation in Tennessee: A report of year 1 implementation*. Retrieved from

<http://team-tn.org/wp-content/uploads/2013/08/Year-1-Evaluation-Report-TNDOE.pdf>

Tennessee Department of Education. (2013). *Teacher evaluation in Tennessee; A report of year 2 implementation*. Retrieved from

http://team-tn.org/wp-content/uploads/2013/08/yr_2_tchr_eval_rpt.pdf

Tennessee Department of Education. (2014). *Tennessee Value-Added Assessment System*. Retrieved from <http://www.tn.gov/education/data/TVAAS.shtml>

Tennessee Educator Accelerator Model. (2014). *Teacher evaluation*. Retrieved from <http://team-tn.org/evaluation/teacher-evaluation/>

Tenth Amendment to the U.S. Constitution. (1791). Retrieved from <http://constitution.findlaw.com/amendment10.html>

U.S. Department of Education. (2009). *Race to the top: Executive summary*. Retrieved from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>

University of Florida. (2000a). *Prototype analysis of school effects*. Gainesville, FL: Value-Added Research Consortium.

University of Florida (2000b). *Measuring gains in student achievement: A feasibility study*. Gainesville, FL: Value-Added Research Consortium.

Yettick, H. (2014). Researchers advise caution on value-added models. *Education Week*, 33(32), 10-12.