

Research Progress Report
Cumulative and Residual Effects Of Teachers on
Future Student Academic Achievement
William L. Sanders and June C. Rivers
November 1996

University of Tennessee Value-Added Research and Assessment Center
225 Morgan Hall
P.O. Box 1071
Knoxville, Tennessee 37901-1071

Contents:

Summary of Findings
Introduction
Methodology
Conclusion
References

Summary of Findings

- Differences in student achievement of 50 percentile points were observed as a result of teacher sequence after only three years.
- The effects of teachers on student achievement are both additive and cumulative with little evidence of compensatory effects.
- As teacher effectiveness increases, lower achieving students are the first to benefit. The top quintile of teachers facilitate appropriate to excellent gains for students of all achievement levels.
- Students of different ethnicities respond equivalently within the same quintile of teacher effectiveness.

Cumulative and Residual Effects of Teachers on Future Academic Achievement

William L. Sanders and June C. Rivers¹

INTRODUCTION

The Tennessee Value-Added Assessment System (TVAAS) was designed and has been demonstrated to be an efficient and effective method for determining individual teachers' influence on the rate of academic growth for student populations (Bock & Wolfe, 1996, Sanders, Saxton, & Horn, (in press); Sanders & Horn, 1995). This method requires three key components: a testing process which produces scales that have a strong relationship to the curriculum and which produces measurement that extends above and below grade level, the construction and ongoing expansion of a longitudinal data base, and a statistical process that enables a multivariate, longitudinal analysis to produce unbiased and efficient estimates of the desired effects.

The TVAAS database (approximately 3 million records for Tennessee's entire grade 2-8 student population) provides access to histories of individual student measurements of achievement in mathematics, reading, language arts, science, and social studies available from TCAP achievement test administrations beginning in 1990 and continuing through 1996. The availability of this data affords the unique opportunity to investigate the cumulative effects of teachers on student academic achievement over grade levels. In other words, does the influence of a effectiveness in facilitating academic growth for his/her students continue when these students advance to future grades?

Thus, the purpose of this research report is to present the preliminary results of estimates of cumulative teacher effects in mathematics from grades 3 to 5 using the data from two of Tennessee's larger metropolitan systems. This research is ongoing and will be expanded to cover a greater diversity of districts, grade levels, and academic subjects. A secondary objective was to decompose the data from teacher effectiveness groups in an attempt to understand which achievement levels of students were being offered opportunities to make satisfactory academic growth. The data were further decomposed to observe any differential responses over ethnic groups.

Methodology

Phase 1 Analysis

The specific data used in this study were restricted to the cohort of students who were second graders in 1991-92, third graders in 1992-93, fourth graders in 1993-94, and fifth graders in 1994-95. Using this data, teacher effects were estimated from a longitudinal analysis by using a statistical mixed model process that provided shrinkage estimation for the teacher effects.

The specific model fitted to the data was:

$$\text{Current score} = a + b * (\text{previous math score}) + t(i) + \text{error}$$

Where a = constant to be estimated from the data

b = regression coefficient

t(i) = shrinkage estimates of the teacher effects.

After the teacher effects were obtained for each grade level, the distribution of teachers was arbitrarily grouped into five quintiles, with the teachers demonstrating the lowest degree of effectiveness in the first quintile and the teachers demonstrating the greatest degree of effectiveness in the fifth quintile. This process was repeated independently for grades three, four, and five for both systems. For the purposes of this specific research, this rather simple model was considered to be adequate to identify groups of teachers within each of the quintiles².

By encoding individual student records with the teacher effectiveness quintiles for each grade, the progress of individual students was traceable through identified sequences of teacher effectiveness. Thus it was possible to determine whether teachers from previous grades affected current year scores.

Phase 2 Analysis

The data from each of the two systems were analyzed independently for one cohort group. Each cohort group analysis encompassed three years of student TCAP achievement scale scores. The specific model for these secondary analyses was:

$$\text{Fifth grade score} = a + b * (\text{second grade score}) + tq3(i) + tq4(j) + tq5(k) + \text{error}$$

Where

a = constant estimated from the data

b = regression coefficient

tq3(i) = quintile of the third grade teacher

tq4(j) = quintile of the fourth grade teacher

tq5(k) = quintile of the fifth grade teacher

Second grade scores were included in the model to insure that the estimates for the subsequent teacher quintiles would not be biased for any disproportionate assignment of students to the various teacher sequences. Models for preliminary analyses contained the interactions of the classification variables, in nearly all cases, these effects were not significantly different from zero, or had only a very small effect on the score. Thus, these variables were excluded from the final model.

Results

From the phase 2 analyses, the F-statistics denoting the residual effects of the third grade teacher quintiles on fifth grade math scores were 16.25 and 14.03 for systems A and B, respectively; for residual effects of fourth grade teacher quintiles, 11.51 and 18.87; and for the direct effects of fifth grade teachers on fifth grade scores, 97.63 and 92.04. All of these effects were very highly significant.

Five quintiles for each of three grades provide for 125 possible teacher-sequence combinations. To denote the magnitude of the cumulative effects of these sequences, estimable functions corresponding to seven of these combinations were chosen and evaluated from the solution vector of the final model. The estimated means and their corresponding percentiles are presented in [figure 1](#).

The difference in fifth grade math achievement means between the *High-High-High* and the *Low-Low-Low* effectiveness sequences is dramatic yet rather consistent for both systems. With second grade scores equalized, in system A, the *Low-Low-Low* sequence resulted in a mean of 720.2 (44th percentile³) and the *High-High-High* sequence, a mean of 784.9 (96th percentile). In system B, the *Low-Low-Low* sequence produced a mean of 704.4 (29th percentile), and the *High-High-High* sequence, 758.9 (83rd percentile). **With an even start, the difference in these two extreme sequences resulted in a range of mean student percentiles in grade five of 52 to 54 points!!**

However, other important differences can be observed in [Figure 1](#). Observe the variance in the comparison of the *Avg-Avg-Avg* and the *High-High-High* sequences: student performance varies from the 79th to the 96th percentile for system A, and from the 50th to the 83rd for system B. The ranges of percentile scores for systems A and B in a similar comparison for the *Low-Low-Low* and *Avg-Avg-Avg* sequences are the 44th-79th and 29th-50th percentiles, respectively. By looking at sequences in which the fifth grade teachers were comparable in terms of effectiveness, it is

possible to see the residual effects of prior year teachers. This type of comparison equalizes the direct effects of the fifth grade teachers on the student achievement scores so that the variability is attributable to the prior year combinations (Note: the analysis also equalized second grade scores). A comparison of the *Low-Low-High* with the *High-High-High* sequences demonstrates the extreme of residual teacher effects on student performance. This comparison shows a difference of thirteen percentile points for system A and twenty-four percentile points for system B. A further comparison of the Low-Low-High and Avg-Avg-High sequences provides a less dramatic example, but a significant one, nonetheless. In system A, two years of ineffective teachers versus two years of moderately effective teachers produced a difference of nine percentile points in student performance. In system B, the same comparison shows a difference of eleven percentile points.

As was mentioned previously, there were not important interactions between the teacher quintile groups over grades. This absence of interaction implies that the teacher effects are cumulative and additive with very little, if any, suggestion of compensatory effects. **An effective teacher receiving students from a relatively ineffective teacher can facilitate excellent academic gain for his/her students during the school year. Yet these analyses suggest that the residual effects of relatively ineffective teachers from prior years can be measured in subsequent student achievement scores.**

Phase 3 Analysis

The effects of teacher sequences on student achievement, as presented above, provide information relative to group averages, yet these analyses do not provide information as to the effectiveness for specific achievement levels of students best or least served by the different levels of teacher effectiveness. In [Table 1](#), student gains, averaged by achievement level of the students⁴, were cross tabulated with teacher quintile groups and are presented for both school systems. This presentation is restricted to fifth grade student achievement.

Results

The target gain for fifth grade math achievement is an average of 25 scale score points. In [Table 1](#), a comparison of average student achievement gains with this target gain shows the first quintile of teachers to be ineffective with *all* achievement levels of students. The second quintile of teachers facilitated this degree of achievement with the lower achieving group, but became less effective as the achievement level of the students increased. Although the third quintile of teachers was effective with more achievement levels, lower achieving students profited more than higher achieving students when assigned to "average" teachers in both systems. Teachers in the fourth quintile achieved target gains with all but the highest level of student achievers; and again, the lower achieving students were better served. The fifth quintile teachers were generally effective with **ALL** student achievement levels, but even the highest achieving students made less than adequate gains in one of the two systems. **In both systems, teachers in the two lower quintiles did not facilitate target gains with most of their students; and overall, a greater percentage of low achieving students than high achieving students made satisfactory gains.**

Phase 4 Analysis

A common concern of child advocate groups is the potential of disproportionate assignment of minority students to inadequate teachers. Bridges (1996), in a review of teacher evaluation and ensuing personnel assignment practices, substantiated this concern. Bridges found that when parents and students complained about inadequate teachers, in many instances the inadequate teachers were transferred to schools where no one was likely to complain about their performance. Typically, the teachers were transferred to schools with one or more of the following characteristics: schools with high student transfer rates, schools with large numbers of students receiving free or reduced price meals, schools with high numbers of minority students, schools with high numbers of students who were considered to be "disadvantaged" in some way by the educational community. The final analysis of this report focuses on the relationship of teacher

effectiveness and ethnicity across student achievement levels. The purpose of the analysis was to ascertain whether teacher effectiveness varies between ethnic groups of students. This analysis was limited to the differentiation of black and white third grade students from system B and their teachers. System B was chosen because of the greater ethnic diversity within the student population.

Step 1

The number and percentage of black and white students in each quintile was accumulated, and the ethnic group percentage of the total third grade population within the system was calculated. The results of these calculations are presented in [Table 2](#).

Step 2

In an analysis comparable to that of Phase 3, the system B third grade students were further subdivided within student achievement levels into subgroups of black and white students. For a comparison, the average student gain for both black and white students for each achievement level, as well as the average student gain for each quintile of teacher effectiveness, is presented in [Table 3](#).

Results

As shown in [Table 2](#), approximately sixty-two percent (62%) of the third graders in System B were white and thirty-eight percent (38%) were black. Approximately ten percent (10%) more black students than would be expected, based on the ethnic makeup of the system were assigned to the least effective teachers. At the same time, the ratio of white students to black students for the most effective teacher quintile was 3:1; in a distribution of students that paralleled the ethnic makeup of the system, one would expect a ratio of 3:2. The additional ten percent (10%) of black students assigned to the most ineffective teachers represents one tenth of the black third graders in system B.

The target math gain for third grade is 60 points. Looking at [Table 3](#), it appears that, as before in the Phase 3 analysis, the degree of teacher effectiveness is slanted toward lower achieving students in both ethnic groups. For example, the above average teachers of quintile four facilitated average gains in scale score points of 73.3 (for white students) and 74.7 (for black students) in the 600-649 achievement group, but the average gains of this same group of teachers were measured at 55.0 (for white students) and 48.9 (for black students) in the 650-699 achievement group. Yet, the performance of both black and white students within achievement-level subgroups were generally comparable for the teacher effectiveness quintile in most instances. The students within the 600-649 scale score subgroup showed the most consistent comparable gains for the two ethnic groups at every level of teacher effectiveness. **These results suggest that although the student assignment from ethnic groups to effective teachers is slightly disproportionate, the achievement within the two ethnic groups is comparable across the five levels of teacher effectiveness. These analyses suggest that students of the same prior levels of achievement tend to respond similarly to teacher effectiveness levels.**

CONCLUSION

Wright, Horn and Sanders (in press) have demonstrated that, within grade levels, the single most dominant factor affecting student academic gain is teacher effect. The present studies, expanding on the earlier research, strongly suggest the presence of cumulative effects of teachers on student achievement. Groups of students with comparable abilities and initial achievement levels may have vastly different academic outcomes as a result of the sequence of teachers to which they are assigned. These analyses also suggest that the teacher effects are both additive and cumulative with little evidence of compensatory effects of more effective teachers in later grades. The residual effects of both very effective and ineffective teachers were measurable two years later, regardless of the effectiveness of teachers in later grades.

What can be learned from the academic gain patterns of the varying student achievement levels to refine the characterization and subsequent professional support of relatively effective and ineffective teachers? Regardless of initial achievement level, teachers in the top quintile facilitated desirable academic progress for *all* students. However, regardless of their entering achievement levels, students under the tutelage of teachers in the bottom quintile made unsatisfactory gains. As the teacher effectiveness quintile increased, lower achieving students were first to benefit, followed by average students and, lastly, by students considerably above average.

In terms of student achievement, how can administrators move beyond the numbers of quantitative analyses of this type to improve learning opportunities for *all* students? The results of this study suggest two very distinct opportunities for educational administrators. The first is in the area of student assignment, the second is in the area of formative teacher evaluation in conjunction with preservice and professional development.

Based upon these results, students benefiting from regular yearly assignment to more effective teachers (even if by chance) have an extreme advantage in terms of attaining higher levels of achievement. **(The range of approximately 50 percentile points student mathematics achievement as measured in this study is awesome!!! Differences of this magnitude could determine future assignments of remedial versus accelerated courses.)** In fairness to children of all achievement levels, teacher assignment sequences should be determined to insure that no child is assigned to a teacher sequence that will be unduly hurtful to his or her academic achievement. Of course an administrator's latitude in making student assignments is limited to the existing teacher resources. Even within the context of current teacher resources, administrators should insure that no student is assigned to a very ineffective teacher more than once, and even then insure that each student so assigned, has a highly effective teacher before and after.

The other primary area for improved student achievement is the development and implementation of strategies which will lead to improved teacher effectiveness. As a first step, teachers should be assisted in the use of all available indicators of student academic growth to enable them to identify their own relative strengths and weaknesses. This could include the TVAAS teacher reports, the break-out of class gains by achievement levels⁵, and other formative evaluation tools.

In summary, these results suggest that with appropriate measurements of teacher effectiveness, administrators have undeniable opportunities to minimize the near-permanent retardation of academic achievement of many students resulting from experiencing the most hurtful teacher sequences. If the magnitude of the cumulative effects is not diminished, then students are *de facto* being placed involuntarily in a lottery where the "luck of the draw" of the teacher sequence may play a most important role in their life's opportunities.

REFERENCES

- Bock, R.D. & Wolfe, R. (1996). Audit and review of the Tennessee Value-Added Assessment System (TVAAS): Final Report, Part 1. In Bock, R.D., Wolfe, F. & Fisher, T.H., *A Review and Analysis of the Tennessee Value-Added Assessment System* (Nashville: Tennessee Comptroller of the Treasury, Office of Education Accountability), pp. 69-72.
- Bratton, S.E., Jr., Horn, S.P. & Wright, S.P. (1996). *Using and Interpreting Tennessee's Value-Added Assessment System: A Primer for Teachers and Principals*. (booklet) Knoxville, TN: Univ. of Tennessee, pp. 15-18.
- Bridges, E.M. (1996). Evaluation for tenure and dismissal. In Millman, J. & Darling-Hammond, L. (Eds.), *The New Handbook of Teacher Evaluation* (pp. 147-157). Newbury Park: Sage Publications.
- CTB/McGraw Hill (1990). *Comprehensive Test of Basic Skills, 4th edition, Spring Norms Book*. Monterey, CA: Phi Delta Kappan, 71(6), pp. 104-115.
- Sanders, W.L. & Horn, S.P. (1995). *The Tennessee Value-Added Assessment System (TVAAS):*

Mixed Model Methodology in Educational Assessment. In Shrinkfield, A.J., & Stufflebeam, D. (Eds.), *Teacher Evaluation: Guide to Effective Practice* (pp. 337-350). Boston, MA: Kluwer.

Sanders, W.L., Saxton, A.M. & Horn, S.P. (in press). The Tennessee Value-Added Assessment System (TVAAS): A quantitative, outcomes-based approach to educational assessment. In Millman, J. (Ed.). *Assuring accountability...? Using gains in student learning to evaluate teachers and schools*. Thousand Oaks, CA. Corwin Press.

Wright, S.W., Horn, S.P. & Sanders, W.L. (In submitted manuscript). Teachers and Classroom Heterogeneity: Their Effects on Educational Outcomes. In Iwanicki, E. F. & Schwab, R.L. (Eds.). *Journal of Personnel Evaluation in Education*. Boston: Kluwer Academic Publishers.

Annotations

¹ William L. Sanders, Professor and Director of The University of Tennessee Value-Added Research and Assessment Center (UT-VARAC)

June C. Rivers, Graduate Student, The University of Tennessee, College of Education.

² This model would not be adequate and appropriate to provide the best possible estimate of an individual teacher effect. Rather the full TVAAS teacher model should be used (Sanders, Saxton & Horn, in press). The model employed in this study was fitted to the data via PROC MIXED within the SAS system.

³ Denotes the corresponding percentile (CTB/McGraw Hill, 1990, pp. 104-115).

⁴ Students were classified into achievement subgroups with each subgroup spanning fifty scale score points. The average of each student's previous and current year's mathematics scores were used for this calculation.

⁵ See *Using and Interpreting Tennessee's Value-Added Assessment System: A Primer for Teachers and Principals*, pp. 15-18, by Bratton, Horn & Wright for an explanation of calculating student gains.

[Return](#) 