

**Value-Added Achievement Results for Three Cohorts of  
Roots and Wings Schools in Memphis: 1995-1999 Outcomes**

William L. Sanders and S. Paul Wright  
The University of Tennessee

Steven M. Ross and L. Weiping Wang  
The University of Memphis

July 2000

Value-Added Achievement Results for Three Cohorts of  
Roots and Wings Schools in Memphis: 1995-1999 Outcomes

This special report extends our prior study (Sanders, Wright, & Ross, 1999) of longitudinal student achievement at schools in the Memphis City Schools (MCS) district that implemented the Roots and Wings design. Roots and Wings is a comprehensive elementary school reform design developed by Robert Slavin and his associates at Johns Hopkins University (Slavin, Madden, & Wasik, 1996). The core of the design is Success For All, a research-based program that focuses on developing reading skills using strategies for early intervention and prevention of reading failure. Roots & Wings extends SFA by adding MathWings, a constructivist-based mathematics program for grades 1-5, and WorldLab, a project-oriented curriculum using simulations of real-world events to integrate social studies, science, reading, writing, mathematics, fine arts, and other areas. Additional major components of Roots & Wings are individual tutoring, regrouping of reading classes, a full-time program facilitator, family support, and extensive ongoing professional development for teachers. Roots & Wings was one of the original 11 designs selected nationally for funding by New American Schools (Stringfield, Ross, & Smith, 1996).

### Design

The present analyses examined the progress of 8 Roots and Wings schools that began implementation in the 1995-96 school year, 4 that began in 1996-97, and 10 that began in 1997-98. The time span for the data analyses was spring testing for the years 1994-95 to 1998-99. Thus, in the most recent year of testing (1999), the three cohorts of Roots and Wings schools had respectively completed four, three, and two years of implementation.

These schools were compared in inferential statistical analyses to 23 MCS elementary schools that did not implement a comprehensive school reform (CSR) design and thus, were established as control schools. In prior years, we had employed a matched sample of control schools, but due to changes in schools' status (specifically, control schools choosing to implement CSR designs in different years), the control group could not be maintained. However, because the present study uses "value-added" achievement data from the Tennessee Value-Added Assessment System (TVAAS; see description below), controls for student prior achievement, socioeconomic status, and mobility are built-in to the assessment, making a matched-control school design much less critical. So that readers can gauge the progress of Roots and Wings schools relative to other school groupings, we also provide, in our reporting of descriptive data, the outcomes for other schools implementing CSR designs in Memphis each year and all elementary schools in the State of Tennessee.

#### Achievement Analyses

Student achievement was assessed using "value-added" achievement scores on the Tennessee Value-Added Assessment System (TVAAS) developed by William Sanders and associates at The University of Tennessee. The rationale for TVAAS was to provide student and teacher performance scores free of the biases normally associated with standardized test outcomes data (Sanders & Horn, 1995a, 1995b). By measuring the amount that students gain in their standardized test scale scores from one year to the next, TVAAS scores reflect growth regardless of initial level of performance. The resultant estimates of student and teacher effects have been demonstrated to be statistically independent of socioeconomic confounding and do not require direct measures of these variables (Sanders & Horn, 1995a, 1995b; Wright, Horn, & Sanders, 1997).

The time period for the data analyzed in the present study ranges from spring-1995 to spring-1999. In each of those years, students throughout the State of Tennessee were mandated to participate in the Tennessee Comprehensive Assessment Program (TCAP). From 1995-97, TCAP was a form of the CTBS-4 (CTB/McGraw-Hill, 1990). In 1998, the state-mandated standardized test employed for TCAP was the TerraNova or CTBS-5 (CTB/MacMillan/McGraw Hill, 1997). An equating analysis conducted by William Sanders and his staff in 1999 allowed for the conversion of TCAP scores to the TerraNova scale so that longitudinal effects could be determined (personal communication, William Sanders, April, 1999). In Memphis City Schools, CTBS/5 had been administered in Grade 3, thereby allowing for the computation of TVAAS scores in grades 4 and 5 for five subjects (math, reading, language, science, and social studies) and all subjects averaged.

The index of student achievement used in the present analyses is *the Cumulative Percent of Norm (CPN)* mean. This statistic indicates across all grades reported the percent of the national norm gain. For example, if School A had a CPN gain of 100% in math, it would have achieved at the national or expected level of achievement gain for that subject for that year.

## Results

Because the scores on the five subtests of TerraNova are highly correlated, the results below are reported for all subjects averaged, which is considered a more reliable measure of achievement than any given subtest and substantially less subject to year-to-year fluctuations in value-added scores due to item content or difficulty. Descriptive results for the R95, R96, and R97 Roots and Wings cohorts relative to different comparison groups are shown in Figures 1 to 3, respectively.

### 1995 Cohort

Descriptive results. Figure 1 summarizes CPN results (see description above) for the 8 R&W schools that began restructuring in 1995 (“R&W-95”) compared to 12 schools that implemented “Other” restructuring designs, 23 control or non-restructuring (NR) schools, and the 839 elementary schools in the State of Tennessee. The Other designs consisted of ATLAS, Expeditionary Learning Outward Bound, Audrey Cohen College, Modern Red Schoolhouse, Co-NECT, Paideia, and Accelerated Schools. As can be seen from the figure, during the pre-reform year, R&W-95 began by scoring lower ( $M = 109$ ) than the NR schools ( $M = 122$ ) and comparably to the Other designs and the State (both  $M$ 's = 109). In 1996, after one year of implementation, the R&W-95 schools scored similarly to all three comparison groups, including NR (note that the test was more difficult that year for the State in general). In 1997, however, after two years of implementation, R&W-95 schools ( $M = 138$ ) were averaging substantially above the comparisons groups (all approximating  $M = 100$ ), and continued to show comparable superiority in both 1998 (e.g.,  $M$ 's = 123 for R&W and 112 for NR) and 1999 ( $M$ 's = 139 for R&W and 108 for NR). Clearly, in the latter three post-reform years, R&W emerged as highly successful relative to all comparison groups in promoting student achievement gains.

Inferential results. Inferential comparisons, derived from an overall repeated-measures (program x year) ANOVA, compared R&W means to NR means. Results were significant for 1997 ( $t = 3.33$ ,  $p = .001$ ,  $ES = +1.37$ ) and 1999 ( $t = 2.45$ ,  $p = .016$ ,  $ES = +1.03$ ), both showing superior gains for R&W. In 1998, the nonsignificant advantage for R&W did show an  $ES = +0.42$ , a moderately strong effect.

We then compared Roots & Wings-95 to NR for three post-reform years (1997-1999) averaged, as depicted in Figure 4. The R&W post-reform mean ( $M = 133$ ) was significantly superior ( $t = 3.68$ ,  $p = .0004$ ,  $ES = +1.52$ ) to the NR mean ( $M = 107$ ), with comparable advantages relative to Other designs ( $M = 107$ ) and the State ( $M = 105$ ). A related analysis examined the three-year average as a change score relative to the pre-reform (1995) mean. Again, R&W was significantly superior to NR ( $t = 2.76$ ,  $p = .007$ ,  $ES = +1.15$ ).

### 1996 Cohort

Descriptive results. As shown in Figure 2, the 4 R&W-96 schools performed comparably to NR schools and slightly above Other design schools and the State in 1995, but far below these comparison groups in 1996, the spring immediately prior to implementation (e.g.,  $M$ 's = 78 for R&W and 96 for NR,  $ES = -0.61$ ). From 1997 to 1999, R&W-96 consistently surpassed NR and the State (respective  $ES$ s = +0.56, +0.13, +0.14), but to a less striking degree than did the R&W-95 cohort (see above). While R&W-96 cohort also surpassed Other design schools in 1997, it was lower in 1998 and comparable in 1999. As shown in Figure 5, when the post-reform years of 1998 and 1999 are averaged, R&W-1996 ( $M = 114$ ) performed slightly higher than NR ( $M = 110$ ) and the State ( $M = 109$ ), but slightly lower than Other designs ( $M = 118$ ).

Inferential analyses. The inferential analyses comparing R&W-96 to NR were all nonsignificant. However, comparing the pre-minus-post (1995/1996 vs. 1998/1999) change score for R&W-96 vs. that for NR showed a 12.89 point advantage for R&W-96 ( $ES = +0.49$ ).

### 1997 Cohort

Descriptive results. Figure 3 graphically depicts the results for the 10 R&W-97 schools. This cohort performed much lower than the comparison groups in the pre-reform years of 1995 ( $ES = -0.42$ ) and 1996 ( $ES = -0.38$ ), but similarly in 1997 ( $ES = -0.11$ ). In 1998, the first implementation year, they performed above Other designs, but slightly lower than NR ( $ES = -0.19$ ) and State schools. In 1999, however, they performed almost identically to NR ( $ES = -0.07$ ) and the State, but lower than Other designs (whose gain appears to have been artificially boosted to some degree by their exceedingly low 1998 performance).

Inferential analyses. No significant effects were found in comparing R&W -97 schools to NR schools. Comparing the pre-minus-post (1995/1996/1997 vs. 1999) change score for R&W-97 vs. that for NR showed a 6.87 point advantage for R&W-97 ( $ES = +0.20$ ).

### Cumulative Results

In a culminating analysis, we examined how the 22 R&W schools compared overall to the NR schools. Figure 6 depicts weighted-averages for 1995 and 1999, which are common pre- and post-reform years for all cohorts. As the figure shows, R&W schools were lower performing in pre-reform and higher performing in 1999 (two to four years post-reform). We performed two types of inferential analyses. One analysis compared R&W versus NR schools in change score from the last pre-reform year (e.g., 1995 for R95 schools) to the first post-reform year (e.g., 1997 for R95 schools). This analyses yielded a mean difference 22.10 *CPN* points, favoring the R&W schools over the NR schools ( $t = 2.52, p = .013$ ). The second analysis performed parallel comparisons but for all of the school's pre-reform and post-reform years (e.g., 1995 vs. 1997, 1998, and 1999 averaged for the R95

cohort). This outcome was also significant, showing an average advantage of 19.50 *CPN* points, favoring R&W over NR ( $t = 2.16, p = .033$ ).

### Discussion

Student performance was assessed for R&W schools in Memphis, Tennessee relative to control schools in grades 4-5 on five subjects on the TerraNova, the state-mandated standardized achievement test. Value-added scores, derived from the Tennessee-Value Added Assessment (TVAAS) system, were used to assess student gain from one year to the next. In the present analyses, we examined results on all five subjects averaged for three cohorts of R&W schools (R95, R96, and R97), the first of which began implementation in the fall, 1995.

The five-year (1995-1999) TVAAS findings are clearly supportive of R&W's effects in increasing student gains in academic achievement. Taking the results for all 22 R&W schools in the R95, R96, and R97 cohorts combined shows an average pre- to post-reform gain of about 20 points higher than that realized by the control schools. This advantage suggests that the average R&W student was gaining from year to year 20% more of the national norm gain for his/her grade than were the control school counterparts. Such an advantage is not only statistically significant but highly educationally important.

Our results also suggest a trend that is both logical and predictable. Specifically, the positive impacts of R&W in raising student achievement become increasingly evident as the maturity of the program implementation increases (e.g., Slavin et al., 1996). The R95 cohort, for example, did not realize gains until 1997—the second year of implementation. Importantly, these gains continued to be strong in the third and fourth year. Effect sizes in 1997, 1999, and for the three-year (1997-99) average all exceeded +1.00, indicating an

educationally important impact. R96 schools performed poorly in their first year of R&W but higher than control schools thereafter. R97 remained below control schools in all years, but essentially bridged the gap in 1999, following two years of implementation. Successful reform takes time, and a longitudinal research plan that compares results over multiple years is essential for judging the effects of comprehensive models, such as Roots & Wings.

Judging from the present five-year study, Roots and Wings is not only producing gains higher than those at non-restructuring schools, but has emerged, in Memphis, as one of the most successful and consistent designs in raising student achievement.

## References

- CTB/McGraw-Hill (1990). *Comprehensive test of basic skills (4<sup>th</sup> ed.)*. Spring Norms Book. Monterey, CA: CTB/MacMillan/McGraw-Hill.
- CTB/MacMillan/McGraw-Hill (1997). *TerraNova Technical Bulletin I*. Monterey, CA: CTB/McGraw-Hill.
- Sanders, W. L., & Horn, S. P. (1995a). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Educational Policy Analysis Archives*, 3(6).
- Sanders, W. L., & Horn, S. P. (1995b). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. In A. J. Shinkfield & D. Stufflebeam (Eds.), *Teacher evaluation: Guide to effective practice* (pp. 337-350). Boston: Kluwer.
- Sanders, W. L., Wright, S. P., & Ross, S. M. (1999). Value-added achievement results for two cohorts of Roots and Wings schools in Memphis: 1995-1998 outcomes. Memphis, TN: Center for Research in Educational Policy, The University of Memphis.
- Slavin, R. E., Madden, N. A., & Wasik, B. A. (1996). Roots and Wings. In S. Stringfield, S. Ross, & L. Smith (Eds.), *Bold plans for school restructuring: The New American Schools Designs*. Hillsdale, NJ: Erlbaum.
- Slavin, R.E., Madden, N. A., Dolan, L., Wasik, B. A., Ross, S. M., Smith, L. J., & Dianda, M. (1996). Success for All: Summary of research. Journal of Education for Students Placed at Risk, 1, 41-76.

Stringfield, S.C., Ross, S.M., & Smith, L.J. (1996). *Bold plans for school restructuring: The New American Schools designs*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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, 57-67.

Figure 1. Memphis City Schools TVAAS Results for Roots & Wings 1995 vs. Other R95 Designs vs. NR and State for 1995 – 1999.

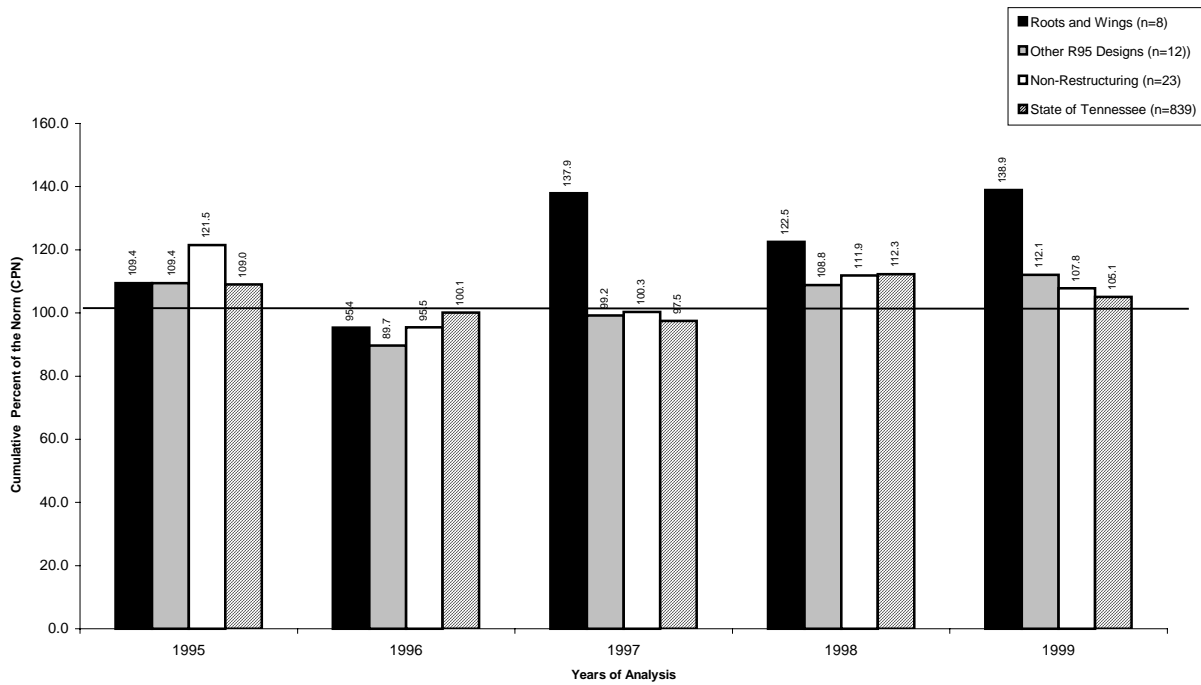


Figure 2. Memphis City Schools TVAAS Results for Roots & Wings 1996 vs. Other R96 Designs vs. NR and State for 1995 – 1999.

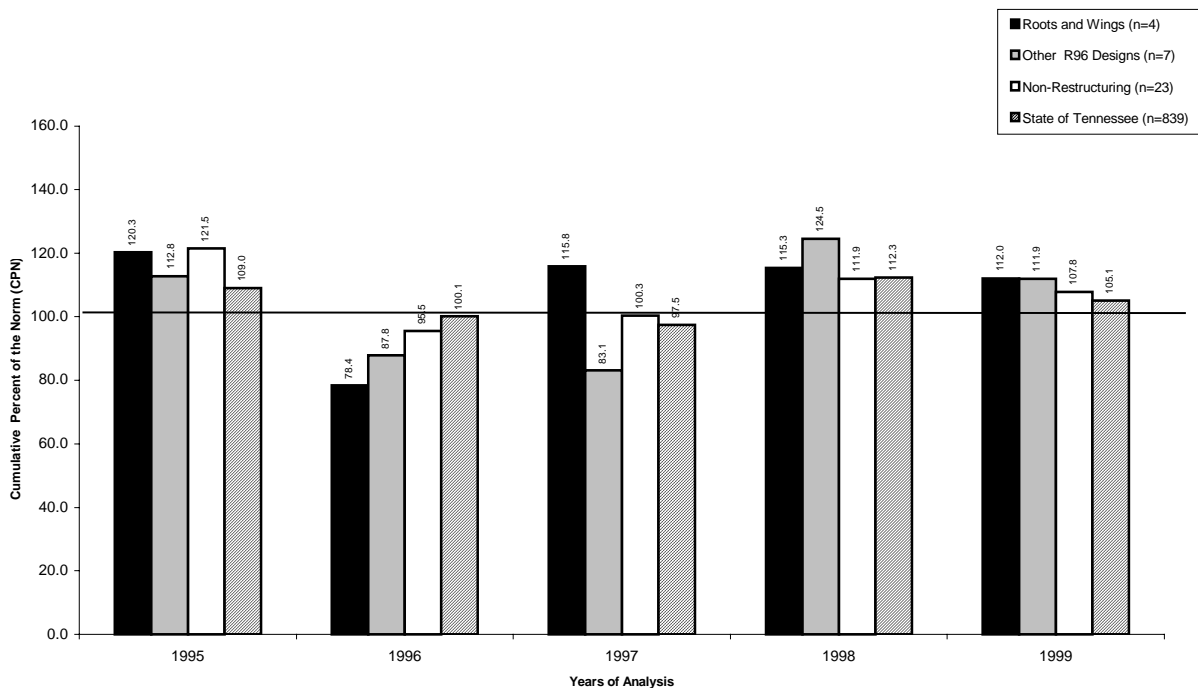


Figure 3. Memphis City Schools TVAAS Results for Roots & Wings 1997 vs. Other R97 Designs vs. NR and State for 1995 – 1999.

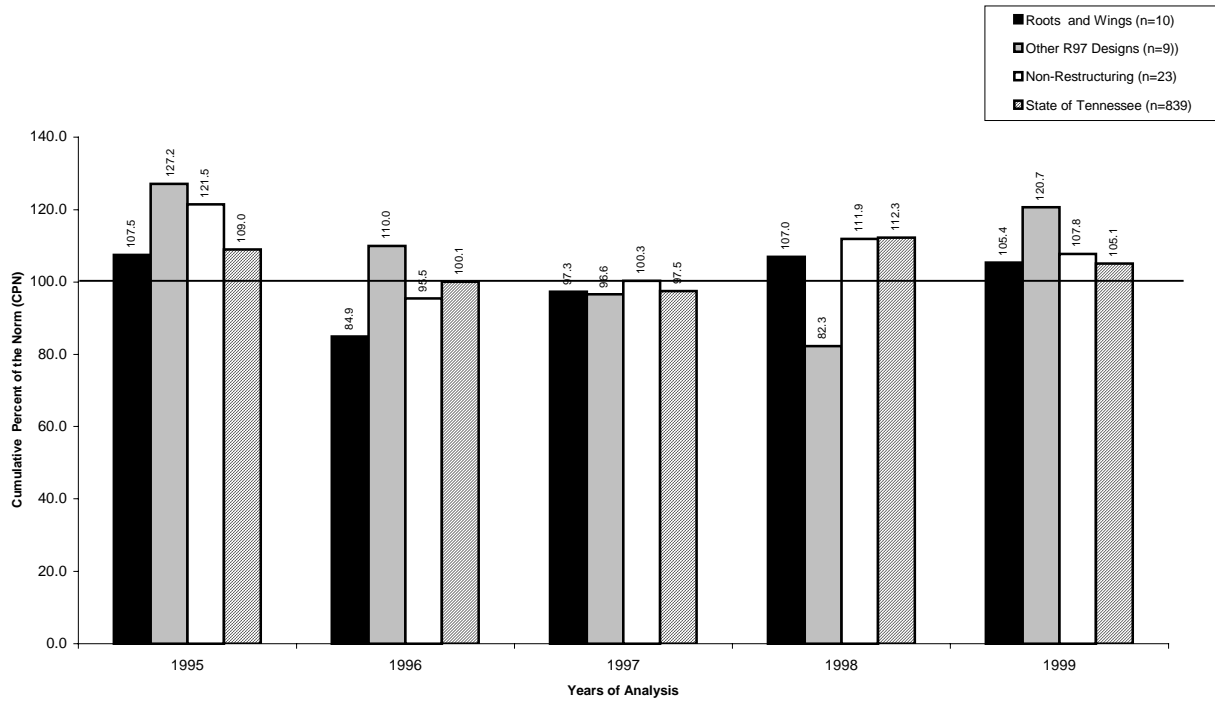


Figure 4. Memphis City Schools TVAAS Results for Roots & Wings 1995 vs. Other R95 Designs vs. NR and State for 1995 and 1997 – 1999.

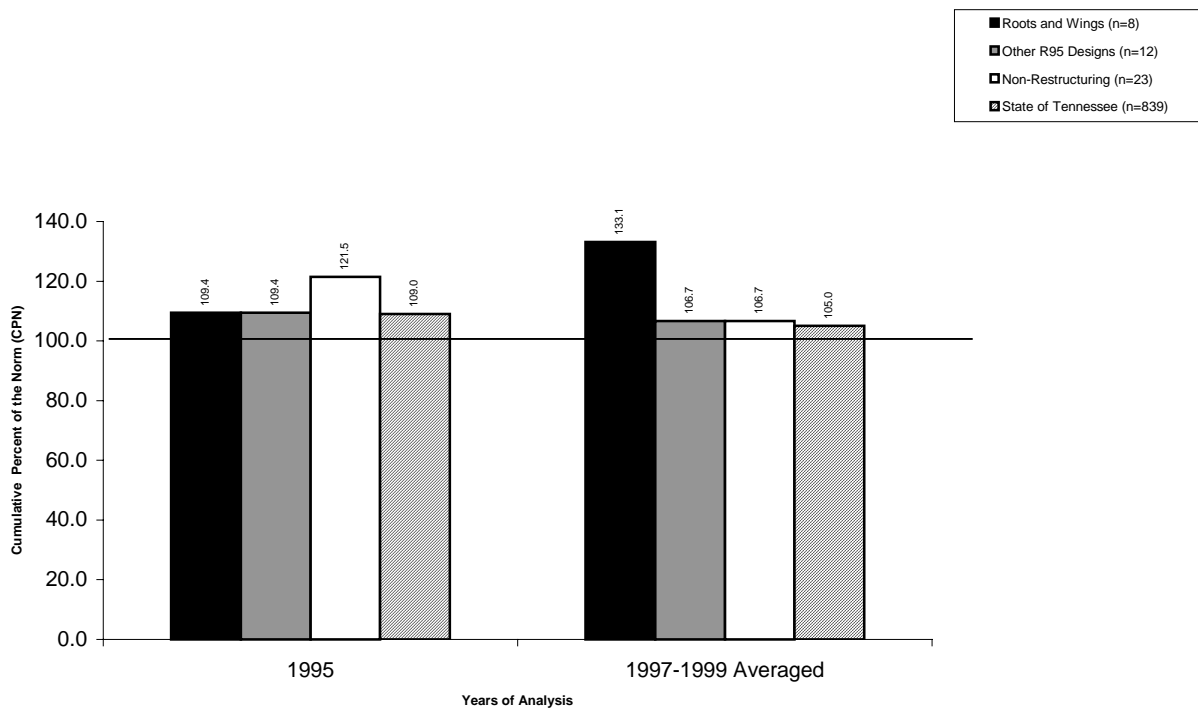


Figure 5. Memphis City Schools TVAAS Results for Roots & Wings 1996 vs. Other R96 Designs vs. NR and State for 1995, 1996 and 1998 – 1999 Averaged.

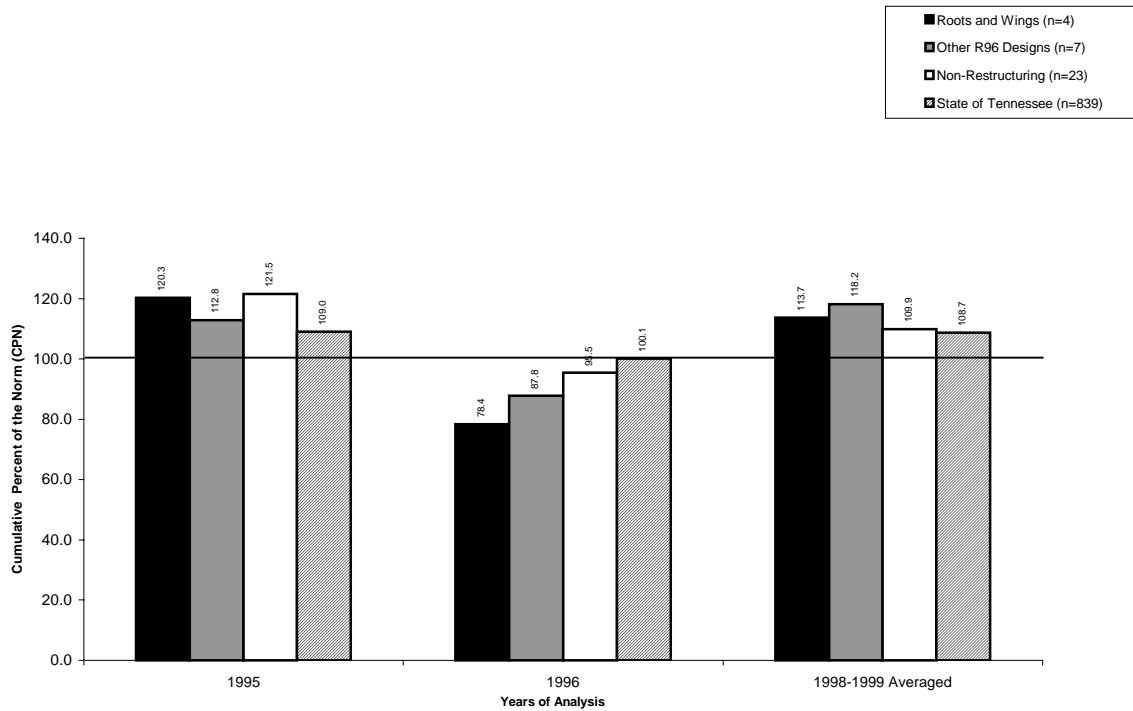


Figure 6. Memphis City Schools TVAAS Results for Roots & Wings 1995 – 1997 Averaged vs. Other R95 – R97 Designs vs. NR and State for 1995 and 1999.

