POLICY STUDIES ASSOCIATES, INC.

ACHIEVEMENT TRENDS IN SCHOOLS WITH SCHOOL ADMINISTRATION MANAGERS (SAMs)

Brenda J. Turnbull
Richard N. White
Erikson R. Arcaira

August 2010

Policy Studies Associates, Inc.
1718 Connecticut Avenue, NW
Suite 400
Washington, DC 20009
www.policystudies.com

For:
The Wallace Foundation
5 Penn Plaza
New York, NY 10001

Summary

The School Administration Manager (SAM) project, supported by The Wallace Foundation as part of its work on educational leadership, focuses on changing the conditions in schools that prevent principals from devoting more time to instructional leadership. In schools participating in the National SAM Project, principals have made a commitment to increase the amount of time they spent on tasks related to instructional leadership, and schools have hired or designated a school administration manager (called a SAM) to support and assist the principal in changing his or her practice. Often, principals have delegated administrative and managerial tasks to SAMs. Participating teams have also received software for recording and analyzing the principal's time use, and monthly coaching from experienced administrators.

Since 2008, Policy Studies Associates (PSA) has worked with The Wallace Foundation to evaluate the implementation, diffusion, and early effects of the SAM project. The 2009 report from the evaluation examined program effects on principal time use and instructional leadership.[1] This follow-up report describes student achievement trends in schools participating in the SAM initiative and groups of similar comparison schools. In addition to the overall comparison between participating and nonparticipating schools, the report also analyzes student achievement trends in the subset of participating schools where the principals made the greatest changes in their use of time.

By the end of 2009 PSA could examine data from five cohorts of SAM schools, a total of 54 SAM schools in Kentucky, Illinois, and Iowa in which students had been tested at least twice after introduction of the SAM project. Evaluators used a matching process to select comparison schools and compare achievement trends. Overall, the analysis of school-level achievement data found a mixed picture on student achievement. The findings are summarized here:

- Each of the five cohorts of participating schools had, on average, an increase in achievement in the *first year* following introduction of the SAM project
- These one-year increases were not significantly different from the achievement increases observed in matched comparison schools that did not participate in the SAM project
- Four of the cohorts (22 of the 54 SAM schools in total) had increases in achievement over the *two years* following introduction of the SAM project
- In two cohorts (13 of the 54 SAM schools in total) the two-year increases in achievement exceeded those of their matched comparison schools, and these differences were statistically significant

[1]Turnbull, Brenda J., Haslam, M. Bruce, Arcaira, Erikson R., Riley, Derek L., Sinclair, Beth, & Coleman, Stephen, "Evaluation of the School Administration Manager Project," December 2009. Available at: .http://www.wallacefoundation.org/KnowledgeCenter/KnowledgeTopics/CurrentAreasofFocus/EducationLeadership/Pages/evaluation-of-the-school-administration-manager-project.aspx

The principals' time use in these schools, as measured by the SAM project after one year of participation, did not have a statistically significant relationship with student achievement gains. Participating principals varied in the extent to which they increased their time spent on instruction and in the absolute percentage of their time they spent on instruction, but neither the increase nor the percentage was correlated with student achievement gains. Moreover, those schools in which principals met the project's short-term goal for change in time use generally did not outpace their matched comparison schools (schools not participating in the SAM project) in student achievement gains. There was an exception: in one cohort, the four schools whose principals met this time-use goal had significantly higher achievement gains than their matched comparison schools after one year; this was no longer true after two years, however.

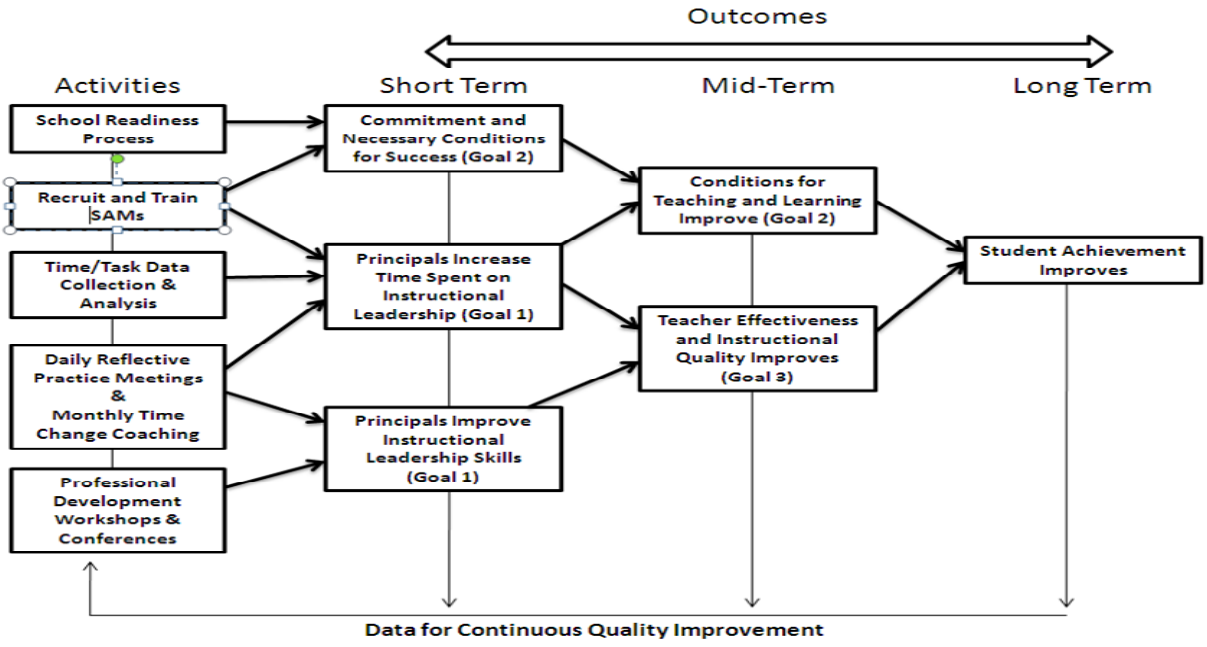
Introduction: The SAM Project

As part of its multifaceted work in educational leadership, The Wallace Foundation supported the development and expansion of the School Administration Manager (SAM) project. In this project, participating principals' time use is measured by trained data collectors at the start of participation, and each school hires or designates a SAM to continue tracking time use and assist the principal in changing his or her practice. SAMs are expected to meet daily with the principal to discuss his or her time use; many also take on management tasks in the school.

The SAM project provides tools, coaching, and data analysis to help principals change their practice. Principals are asked to record their activities in detail during the day using the TimeTrack system, which provides a means for principals to reflect on their practice by coding time use by task and category (instruction, management, and personal time) and by producing weekly and monthly time use reports. The project also requires principal/SAM teams to meet monthly with a Time Change Coach, typically a retired school administrator, who is selected and trained to discuss progress and identify training needs with the team. At 12-month intervals, a trained data collector returns to record the principal's behavior at five-minute intervals, generating a Time/Task Analysis report. These reports provide an external assessment of time use for individual principals and, in the aggregate, for the National SAM Project.

The Wallace Foundation supported development of a logic model for the project (Exhibit 1). The model shows that these various activities and supports were designed to change principals' time use and thereby to improve organizational health and instructional quality and, ultimately, raise student achievement. Compared with an earlier logic model, this revised model gives more prominence to the role of professional development in strengthening principals' instructional leadership skills.

Exhibit 1: SAM Project Logic Model



Earlier Evaluation Findings

PSA's 2009 report concluded that the SAM project had effects on principal time use. Specifically, we found:

- Across the 75 principals who had participated in the project for a full year, the time devoted to instruction-related tasks rose significantly over that year. Pre- and post-measures of principals' time use showed that the percent of time devoted to instructional tasks (as defined in detail in the project's system of measurement) rose from a mean of 32 percent to a mean of 45 percent. The change in time use across principals was highly statistically significant, and the effect size was 0.93 of a standard deviation.

- In schools where SAMs carried out five time-consuming management tasks (student discipline, student supervision, managing non-teaching staff, managing school facilities, and interacting with parents), principals' time spent on instruction rose significantly over one year. But where fewer of these tasks were delegated, the principals' time use did not change significantly.

- In terms of actual time spent on instruction, the principals with one-year follow-up data had increased their time by 58 minutes per day or 4 hours 50 minutes per week, on average. The instructional leadership activities on which principals spent at least 10 additional minutes per day were "observation and walkthrough," "instruction-related office work prep," and "work with students." They spent at least 10 fewer minutes per day on each of four management activities: "office work prep," "building management," "student supervision," and "student discipline."

The evaluation findings suggested caution in expecting school improvement to result automatically from changes in principals' time use, however:

- Participating principals welcomed the chance to spend more time in classrooms, and they carried out observations, but our interviews suggested that more could be done to encourage and support principals in deepening their repertoire of leadership skills. Only in a few sites did a principal or others in the school convey a sense that the principal had strategically selected a set of high-leverage leadership activities that would serve specific purposes in school improvement.

Many state and local educators who had adopted the SAM project said that they expected participation in the project to spark school improvements that would be reflected in student achievement gains. They alluded to achievement data reported from the first three schools with SAMs.[2] In those three schools, achievement data from Kentucky's Commonwealth Accountability Testing System (CATS) were examined. Key reported results were:

[2] The report is available from the online Educational Resources Information Center as ED490688. Shellinger, Mark, "Alternative School Administration Study," October 11, 2005, http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/c3/66.pdf

> All three schools increased the rate of academic achievement gain with the School Administration Manager model. All three schools outperformed the district, Jefferson County Public Schools, and the state of Kentucky's rates of gain. (p. 6)

The analyses reported here represent an assessment of the achievement results across these three schools and more than 50 others that have subsequently joined the project. By the end of 2009, PSA could examine data from a total of 54 SAM schools in Kentucky, Illinois, and Iowa in which students had been tested at least twice after introduction of the SAM project (i.e., schools in which SAM project participation began early in the 2007-08 school year, if not earlier, and continued through spring 2009, with school-level achievement results available from 2007, 2008, and 2009). This paper reports on these schools' achievement trends on state tests over time, on comparisons of their achievement trends with those of matched groups of schools.

Data Sources and Analytic Approach

All data presented in this paper come from schools for which we could track two years of student performance after adoption of the SAM project. In some cases there were changes in staffing, with the original principal or SAM leaving and being replaced; however, each of these schools maintained participation in the project. Most of the SAM schools from which we drew data are in Kentucky, where the SAM project has been under way for the longest time. The schools include:

- The 3 pilot schools, all of which are elementary schools
- 6 "expansion" schools that joined the SAM project in fall 2006, all of which are elementary schools
- 32 schools that joined the SAM project in fall 2007, including 24 elementary schools, 4 middle schools, and 4 high schools

In Kentucky, school-level achievement data combine student performance across grades and subjects. Kentucky's accountability system has undergone several changes during the years of program implementation, but analysts have devised ways of calculating achievement trends across years, and we take advantage of these methods here (while noting that not all analysts agree with this choice).

Prior to the 2006-2007 school year, Kentucky used "school accountability index" scores. These included student performance data in reading, mathematics, science, social studies, arts and humanities, practical living/vocational studies, and on-demand writing. In addition, the school accountability index scores incorporated scores from student writing portfolios and national norm-referenced tests in reading, language arts, and math.

For 2007 and 2008, the assessment procedures and the accountability index were revised. Reading and mathematics tests were administered for additional grades to meet the testing requirements of the federal No Child Left Behind law; the tests for on-demand writing and

practical living/vocational studies were moved to different grades; and students were required to take the ACT or its companion exams, ACT EXPLORE or ACT PLAN. The school accountability index formula for 2007 and 2008 incorporated the additional assessments and used new content area weights and new cut scores for student performance. However, Kentucky developed concordance tables that converted the 2007 and 2008 school accountability index scores into "adjusted" scores, allowing the new scores to be compared to pre-2007 data.

The Kentucky Department of Education's Office of Assessment and Accountability in its presentation[3] of concordance tables to parents and community members explained that:

- *A concordance table allows us to compare two different tests from two different years.*
- *With concordance, we can link the 2007 scores to original school goals.*
- *By doing this, we can keep school accountability going.*

In its release[4] of the 2007 spring test results, the Kentucky Department of Education wrote:

> *"The Adjusted Accountability Index continues accountability. It links the 2007 performance to the past system. The score is related to the performance trends, the original goal and assistance lines in a school's or district's growth chart."*

We used concordance tables to convert the 2007 and 2008 accountability index scores into "adjusted scores" based on the department's description of their comparability.

In 2009, Kentucky's legislature passed Senate Bill 1, which called for the creation of a new system of assessment and accountability and effectively suspended accountability until the 2011-2012 school year. State assessments in mathematics, reading, science, social studies, and on-demand writing were still administered; however writing portfolios were eliminated and arts/humanities and practical living/vocational studies were not tested in 2009. The state did not calculate 2009 school accountability index scores. However, the Prichard Committee and its partners developed "transition index" scores as a means of assessing and comparing school progress. The transition index system integrates performance on five subjects—mathematics, reading, science, social studies, and on-demand writing—and categorizes students into eight different performance levels, with each performance level given a corresponding weight value. Transition index scores were calculated for 2009 and the two previous years (2007 and 2008) but not for earlier years. Transition index scores are not directly comparable to the pre-2007 or adjusted index score system. Consequently, the 2009 progress of the pilot schools and the 2006 SAM cohorts is not presented in this report.

For this report, we analyzed achievement trends up through 2008 using the pre-2007 accountability index scores and the 2007 and 2008 "adjusted" index scores. These scores were used to trace the achievement of the three pilot schools and the cohort of six schools that began

[3] Kentucky Department of Education Office of Assessment and Accountability. "Concordance Table Presentation for Board and Community Members." September 10, 2007.
[4] Kentucky Department of Education. "Kentucky Core Content Test Scores Released." October 2, 2007. http://www.education.ky.gov/kde/homepagerepository/news+room/press+release+and+advisory+archives/2007+press+releases+and+advisories/07-081.htm

implementing the SAM project in 2006. To assess the one-year and two-year gains of schools that began implementation in 2007, we relied on the Prichard Committee's 2007, 2008, and 2009 transition index scores.

Reviewing a draft version of this report, Kentucky assessment experts questioned our use of the concordance tables and of the transition index for analyses over time.[5] Their current stance is that year-to-year comparisons are valid through 2006 and then for 2007-08. Readers may therefore prefer to disregard the comparisons in this report that bridge the years from 2006 to 2007 and beyond, and those that include the use of 2009 data. Our use of these comparisons does not change the report's conclusions; it adds to the number of comparisons made but does not bias the results in one direction or the other.

Illinois SAM schools discussed here include a single cohort:

- The 5 high schools that joined the SAM project before December 2007 and for which student achievement data were publicly available from spring 2007, 2008, and 2009

The Illinois achievement data consist of Prairie State Achievement Examination (PSAE) scores for each school. PSAE scores measure the achievement of 11th grade students in reading, mathematics, science, and writing and incorporate performance on the ACT along with the state-developed assessments.

Iowa SAM schools discussed here include a single cohort:

- The 8 schools that joined the SAM project before December 2007, including 4 elementary schools, 2 middle schools, and 2 high schools

The Iowa achievement data are the average of the proportion of students scoring proficient or above in reading and in mathematics at each school. The proportion is computed as the simple average of the proportion proficient across all tested grade levels served by each school, and across both subjects.

In order to examine achievement outcomes over time, we employed a quasi-experimental design that compares the achievement trends in participating schools with achievement trends in groups of similar but nonparticipating schools. Many factors are statistically associated with school-level academic achievement gains, notably past performance and student demographics. Therefore, to measure the impact of program participation on academic achievement, analyses

[5] Letter to Jody Spiro, The Wallace Foundation, from Mark Shellinger on behalf of the National SAM Innovation Project Program Board, June 17, 2009. With respect to the concordance tables, the letter says that because the tests changed in 2007, "you can't compare index scores pre-2007 with post-2006" (p.2). With respect to the Prichard Committee's conversion of scores, it says, "This conversion is not recognized/used by the Kentucky Department of Education and Jefferson County Schools." Although we believe that the use of both conversion approaches has value in preliminary analyses like those presented in this report, we note that a more conservative approach would limit the analysis to achievement trends through the years ending in 2006 and then for the period 2007-08, would use transition index scores rather than adjusted accountability scores for 2007 and 2008, and would not report 2009 scores.

must compare participating and nonparticipating schools that are similar with respect to these characteristics. For each SAM school, either one or two comparison schools were selected using publicly available data for the year prior to SAM implementation in each school. Using a sampling procedure known as propensity matching, we selected comparison groups of schools for each cohort of SAM schools, drawing samples that matched the SAM schools as closely as possible with respect to size, grade levels served, race/ethnicity, performance on state assessments, free and reduced-price lunch participation, percent of students in special education, and percent of students who are English language learners.

To improve comparability between participating and nonparticipating schools, we applied adjustments to the matching process. For SAM schools in large urban districts, comparison schools were drawn from the same district; for SAM schools in other settings, the comparison schools were drawn from a statewide dataset.

- For SAM schools in Jefferson County, Kentucky, the comparison schools were selected from within Jefferson County[6]
- For other Kentucky SAM schools, comparison schools were selected from districts similar in urbanicity to the SAM school (i.e., city, suburban, town, rural).
- In Illinois, comparison schools were selected from the same district as the SAM schools
- In Iowa, comparison schools were selected from districts similar in urbanicity to each SAM school (i.e., city, suburban, town, rural)

In addition, weights were applied to different variables in the matching process. Among the potential comparison schools meeting the geographic criteria, schools were matched first on achievement, using the achievement measures described above for each state; then on the proportion of students who were non-white, the proportion in poverty, and the proportion eligible for English language learner and special education services; the grade levels served by the school; and total school enrollment. T-tests were conducted to check baseline equivalence on matching variables. These tests revealed no statistically significant differences ($p<.05$) between participating schools and comparison schools within any of the five analysis groups.

In order to analyze the impact of program participation on academic achievement, SAM schools are compared with their matches using paired samples t-tests with a two-tailed test of significance ($p<.05$). Because measures of achievement are not directly comparable across states (or, in Kentucky, across cohorts), the findings are specific to each of the five cohorts of schools joining the project: three cohorts in Kentucky; one in Illinois; and one in Iowa.

[6] Because of the limited number of schools available in this comparison pool, only one matched comparison school was selected for each SAM school in Jefferson County. Then, for consistency of procedures across Kentucky, we also selected only one comparison school for the other Kentucky schools.

Limitations of This Analysis

There are several important caveats for this analysis. The number of schools with sufficient data for multi-year analysis, while larger than in previous years, remained quite small. By opting for measures that combined results across subjects and grades, we sacrificed sensitivity to some changes that might be occurring within schools. Similarly, analyses based on school-level data are far less sensitive to achievement differences than analyses using student-level data. Obtaining and using student-level data would have revealed changes in achievement among the same individual students over time rather than this year's students in the tested grades versus last year's students in those grades; in addition, it would allow analyses for subgroups of students based on baseline levels of achievement and student characteristics. Perhaps most important, the schools and the principals voluntarily participating in the SAMs project are likely to have differed in systematic ways from comparison groups selected after the fact through statistical procedures.

Findings

We compared achievement trends in a number of ways. For long-term comparisons of SAM schools' achievement trends with their own records and with the state as a whole, we focused on the long-term data available from Kentucky. We also analyzed (and display graphically in this report) the one-year and two-year achievement trends in SAM schools in relation to the comparison groups identified for each state and year. Finally, we explored the possibility that greater changes in principal time use might be associated with more positive achievement results.

Trends over Time in SAM Schools

In Kentucky, the state for which concordance tables furnish achievement data for every year from 2000 through 2008, we can examine long-term achievement histories of the three cohorts of participating schools and the state as a whole (Exhibit 2). The 2005 analysis of SAM school performance said that pilot schools notably accelerated their achievement gains in the first full year of SAM project implementation, and therefore we sought to update and extend this analysis. For each cohort, student achievement did rise in the year following SAM adoption (with the time of adoption marked with a vertical line on the graph), and the rate of increase was somewhat steeper than in previous years.[7]

However, trends in Kentucky do not make a strong case for the achievement results of SAM participation. Looking at the schools' prior achievement trends, we see that the first-year achievement gain was not unprecedented in relation to previous years' gains for any of the cohorts, including the pilot schools. Looking at subsequent years, we see that none of the three cohorts sustained an increased rate of gain beyond the first year of participation in the SAM

[7] The first-year gain from 2006 to 2007 for the six schools that began participation in 2006 should be interpreted with caution because of a change in testing (see note 6 above).

project. For the three pilot schools, there were average achievement declines in their third and fourth years of participation (2006-07 and 2007-08).[8] The six schools that began participation in 2006 had an average achievement decline in their second year of participation (2007-08). (The 2009 achievement data do not appear in Exhibit 2 because they are not comparable with data from 2006 or earlier).

Exhibit 2: Achievement Trends 2000-2008 for Kentucky SAM Cohorts and State

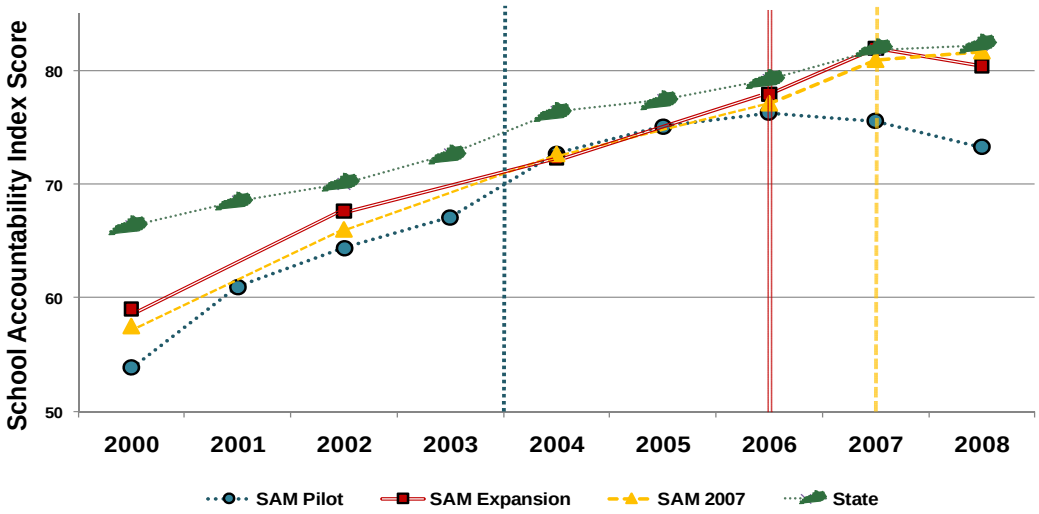


Exhibit Reads: For the SAM pilot schools, the accountability index scores rose unevenly from 2000 through 2003, rose more steeply in the first year of the SAM project (2003-04), then rose less steeply and, beginning in 2006, declined. The SAM expansion schools also had achievement increases in their first year in the project (2006-07); the following year saw a decline in their achievement. Changes from 2006 to 2007 should be interpreted with caution due to changes in testing.

Comparison with Matched Schools

For each cohort of SAM schools, as described above, we identified a group of nonparticipating schools with similar demographics and achievement at the baseline, defined as the year just before participation in the SAM project began. We then analyzed their achievement trends for comparison. The following exhibits show achievement trends in the SAM schools and those of the comparison schools after one and two years.

First, we display the mean one-year change in scores for each cohort of SAM schools and its comparison group (Exhibit 3). In order to give a sense of the magnitude of the gain or loss in achievement, we express these changes in relation to the standard deviation of the baseline scores. Thus it is possible to look across cohorts and see the achievement trends of the SAM schools in relation to those of their comparison schools. The same data are shown graphically in Exhibit 4.

[8]The 2006-07 decline (the third year for this pilot cohort) should be interpreted with caution due to the change in testing (see note 6 above).

Exhibit 3: One-Year Change in Student Performance of SAM and Comparison Schools, By Cohort

Year One		SAM Schools			Comparison Schools		
Cohort	Measure	Baseline Standard Deviation	AverageOne-Year Change in Student Performance	Ratio of Change to Standard Deviation	Baseline Standard Deviation	AverageOne-Year Change in Student Performance	Ratio of Change to Standard Deviation
Kentucky Pilot (n=3 SAM, 3 comparison)	School Accountability Index Scores	17.4	+ 2.3	▲ + 0.13	19.0	- 0.3	▼ + 0.02
Kentucky 2006 Base Year (n=6 SAM, 6 comparison)	School Accountability Index Scores	13.6	+ 3.9	▲ + 0.29	12.1	+ 4.5	▲ + 0.37
Kentucky 2007 Base Year (n=32 SAM32 comparison)	School Transition Index Scores	12.1	+ 1.4	▲ + 0.12	10.8	+ 0.5	▲ + 0.05
Iowa 2007 Base Year (n=8 SAM, 16 comparison)	Average Percent Proficient	10.1	+ 0.3	▲ + 0.03	8.7	- 0.2	▼ - 0.02
Illinois 2007 Base Year (n=5 SAM, 10 comparison)	PSAE Scores	14.4	+ 0.1	▲ + 0.01	14.9	- 3.8	▼ - 0.26

Exhibit Reads: The Kentucky pilot SAM schools, with a baseline standard deviation of 17.4 on their school accountability index scores, made one-year average gains of 2.3, which was 0.13 of a standard deviation; their matched comparison schools had average losses of 0.02 standard deviation over the same year. For the 2006 Base Year group, the one-year gains from 2006 to 2007 should be interpreted with caution due to a change in testing.

Exhibit 4: Standardized One-Year Change in Student Performance for SAM and Comparison Schools, By Cohort

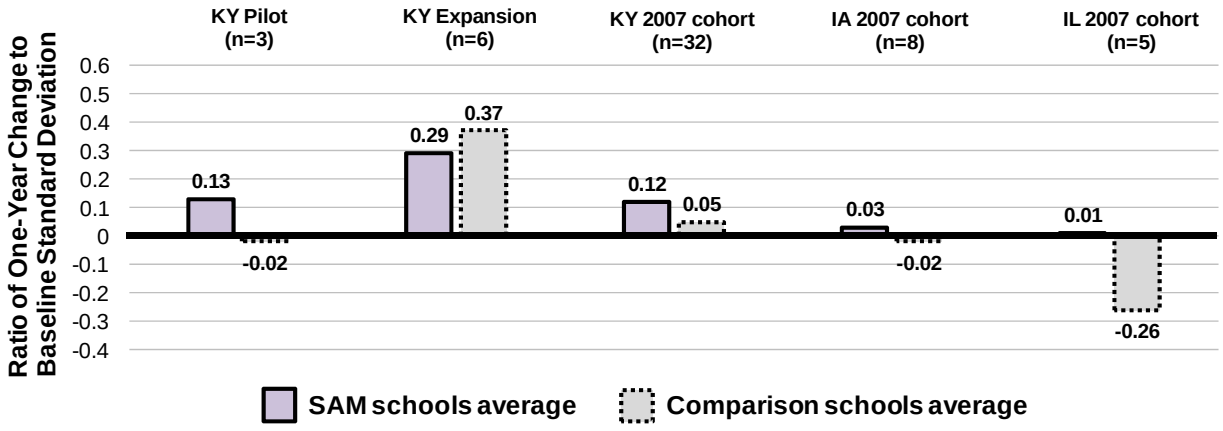


Exhibit Reads: For the three Kentucky pilot schools, the one-year change was an increase of 0.13 of one standard deviation in the baseline scores for that group of schools. The one-year gains from 2006 to 2007 for the "KY expansion" cohort and comparison schools should be interpreted with caution due to a change in testing.

As Exhibits 3 and 4 show, each of the five cohorts of schools had, on average, a one-year increase in achievement following introduction of the SAM project. However, these one-year changes in achievement were not significantly different from the one-year changes in achievement experienced by their matched groups ($p>.05$).

Second, we display the mean two-year change in scores for each cohort of SAM schools and its comparison group (Exhibits 5 and 6). As Exhibit 5 shows, four of the cohorts (22 SAM schools in total) also had an upward achievement trend from baseline to the test administration two years later. However, three cohorts of SAM schools fared no better than the matched comparison groups in achievement over the two-year period.

Exhibit 5: Two-Year Change in Student Performance of SAM and Comparison Schools, By Cohort

Year Two		SAM Schools			Comparison Schools		
Cohort	Measure	Baseline Standard Deviation	Average Two-Year Change in Student Performance	Ratio of Change to Standard Deviation	Baseline Standard Deviation	Average Two-Year Change in Student Performance	Ratio of Change to Standard Deviation
Kentucky Pilot (n=3 SAM, 3 comparison)	School Accountability Index Scores	17.4	+ 3.5	▲ + 0.20	19.0	+ 4.1	▲ + 0.22
Kentucky 2006 Base Year (n=6 SAM, 6 comparison)	School Accountability Index Scores	13.6	+ 2.5	▲ + 0.19	12.1	+ 5.4	▲ + 0.45
Kentucky 2007 Base Year (n=32 SAM, 32 comparison)	School Transition Index Scores	12.1	- 0.2	▼ - 0.02	10.8	- 0.4	▼ - 0.04
Iowa 2007 Base Year (n=8 SAM, 16 comparison)	Average Percent Proficient	10.1	+ 2.2	▲ + 0.22[a]	8.7	+ 0.8	▲ + 0.09
Illinois 2007 Base Year (n=5 SAM, 10 comparison)	PSAE Scores	14.4	+ 0.6	▲ + 0.04[a]	14.9	- 3.0	▼ - 0.20

[a] *Cohort's two-year increases in achievement exceeded those of their matched comparison schools .*

Exhibit Reads: The Kentucky pilot SAM schools, with a baseline standard deviation of 17.4 on their school accountability index scores, made two-year average gains of 3.5, which was 0.20 of a standard deviation; their matched comparison schools had average gains of 0.22 standard deviation over the same period. Note that the two-year trends for the Kentucky 2006 Base Year and Kentucky 2007 Base Year groups should be interpreted with caution, due to changes in testing.

Exhibit 6: Standardized Two-Year Change in Student Performance for SAM and Comparison Schools, By Cohort

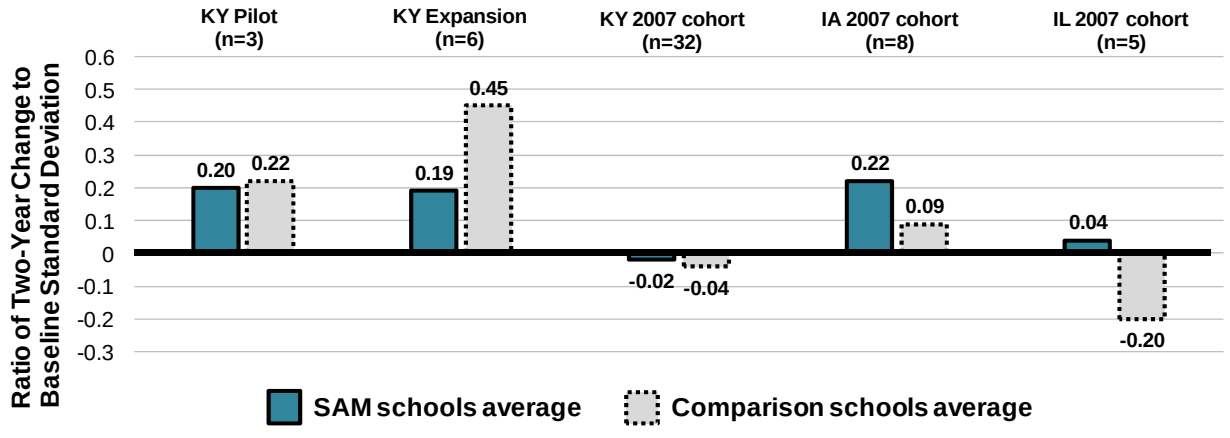


Exhibit Reads: For the three Kentucky pilot schools, the two-year change was an increase of 0.20 of one standard deviation in the baseline scores for that group of schools. Note that the two-year comparisons for the Kentucky Expansion and Kentucky 2007 cohorts should be interpreted with caution, due to changes in testing.

For two cohorts (13 SAM schools in total), the two-year increases in achievement in SAM schools exceeded those of their matched comparison schools: the Iowa and Illinois cohorts had significantly higher achievement gains than their matched comparison schools ($p<.05$).

In response to the concerns expressed by the Kentucky assessment experts, we also analyzed the achievement scores of each cohort using the transition index (rather than "adjusted" accountability scores) for 2007 and 2008. The use of this index made no substantial difference in the findings reported here: trends in scores on the transition index did not greatly favor either the SAM schools or the comparison schools. Similarly, the inclusion or exclusion of 2009 results makes no substantial difference in the report's overall conclusions. All in all, using only a very limited amount of the Kentucky data as suggested by Kentucky reviewers, the Kentucky results continue to show one-year gains in SAM schools that do not significantly exceed gains in the comparison schools, and no difference in achievement results of SAM and comparison schools over a two-year period. (The two-year comparison is made only for the three pilot schools, whose two-year gains from 2004 to 2006 were very slightly less than those of their matched comparison schools.)

Line graphs of achievement over time provide a more detailed look at the two-year achievement trajectories of the SAM school cohorts and their matched comparisons. They give a sense of the magnitude of the achievement fluctuations over time. Here, we return to the use of the concordance tables and adjusted accountability scores for the analysis of Kentucky data. These graphs show that the first two cohorts of SAM schools in Kentucky performed worse than their comparison schools over time (Exhibits 7 and 8) but that the Illinois and Iowa schools that adopted the SAM project in 2007 performed better than their comparison schools (Exhibits 10 and 11). The largest cohort, the Kentucky 2007 cohort, performed very similarly to its comparison group but declined slightly less in achievement (Exhibit 9).

Exhibit 7: Student Performance Trends 2004-2008 for Kentucky Pilot and Comparison Schools

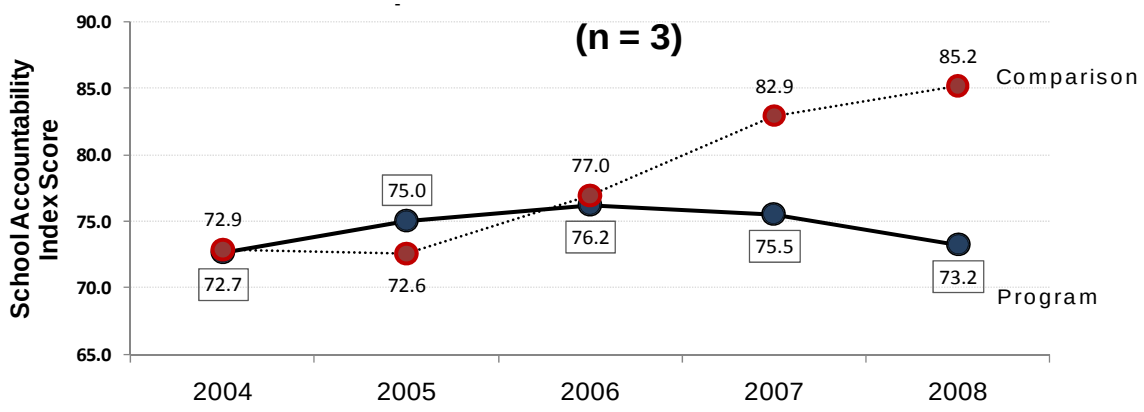


Exhibit Reads: From 2004 to 2005, SAM schools' index scores increased from 72.7 to 75.0 while comparison schools' scores fell from 72.9 to 72.6. Note that changes observed from 2006 to 2007 should be interpreted with caution, due to changes in testing.

Exhibit 8: Student Performance Trends 2006-2008 for Kentucky Expansion and Comparison Schools

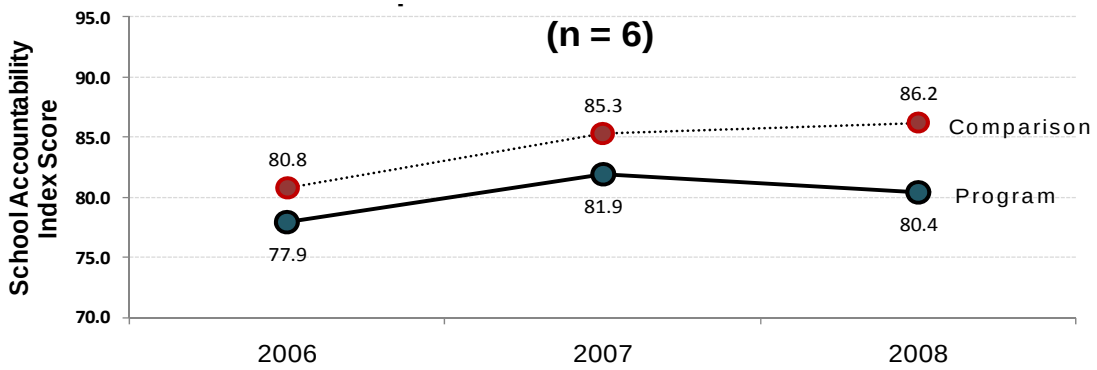


Exhibit Reads: From 2006 to 2007, SAM schools' index scores increased from 77.9 to 81.9. Comparison schools' scores increased from 80.8 to 85.3 during the same one-year period. Note that changes observed from 2006 to 2007 should be interpreted with caution, due to changes in testing.

Exhibit 9: Student Performance Trends 2007-2009 for Kentucky 2007 Base Year and Comparison Schools

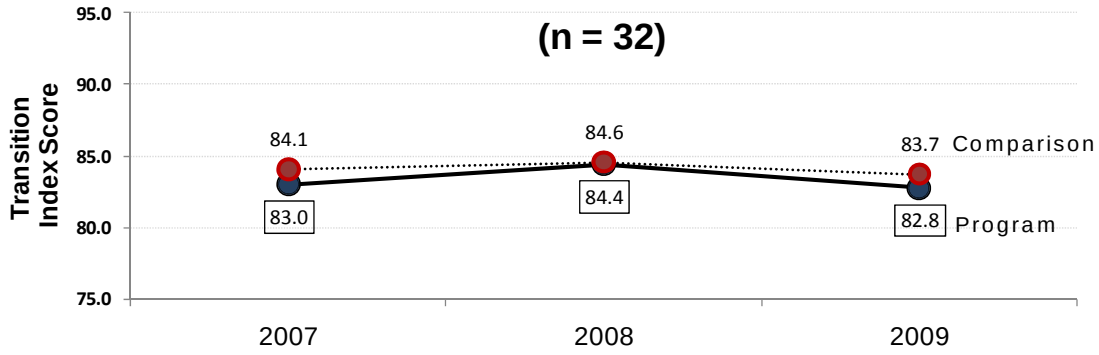


Exhibit Reads: From 2007 to 2008, SAM schools' index scores increased from 83.0 to 84.4. Comparison schools' scores increased from 84.1 to 84.6 during the same one-year period. Note that changes from 2008 to 2009 should be interpreted with caution, due to changes in testing.

Exhibit 10: Student Performance Trends 2007-2009 for Illinois 2007 Base Year and Comparison Schools

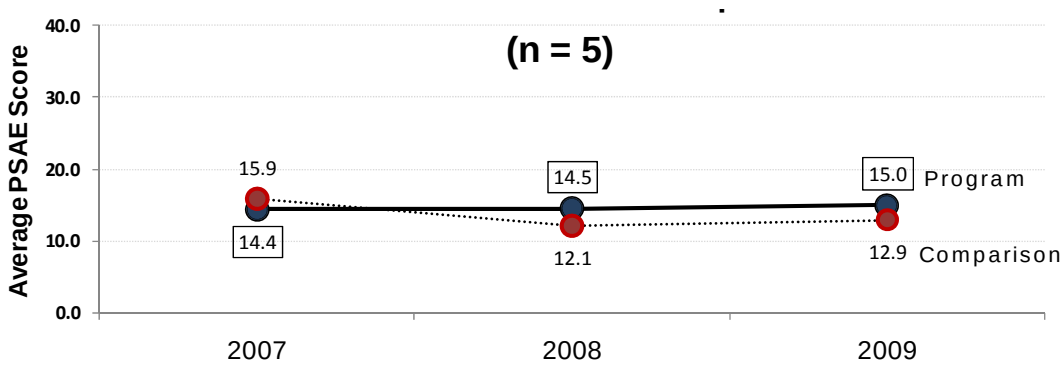


Exhibit Reads: From 2007 to 2008, SAM schools' PSAE scores increased from 14.4 to 14.5 while comparison schools' PSAE scores fell from 15.9 to 12.1.

Exhibit 11: Student Performance Trends 2007-2009 for Iowa 2007 Base Year and Comparison Schools

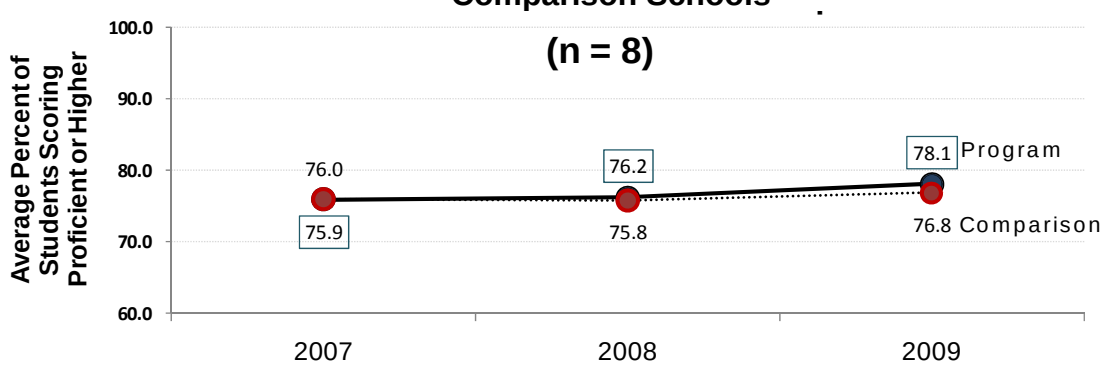


Exhibit Reads: From 2007 to 2008, the average percent of students scoring proficient or higher on their state exams increased from 75.9 to 76.2 in SAM schools while the average percent of students scoring proficient or higher fell from 76.0 to 75.8 in comparison schools.

It is difficult to say whether the achievement results for SAM schools were better than might be obtained by chance. The graphs reveal the fluctuation that can be expected in the average achievement of small groups of schools over a two-year period. It is possible that subgroups of SAM schools, perhaps including most or all of the Illinois and Iowa 2007 cohorts, are implementing the project in a way that makes a systematic difference in student achievement, although the data do not point very clearly to such an inference.

Relationship Between Principal's Time Use and School Achievement

Often, in analyzing achievement in schools participating in an improvement effort, the evaluator looks specifically at the subset of schools that most fully implemented the innovation. The rationale is that these schools' results demonstrate the outcomes attainable when the innovation is implemented as intended.

Since the SAM project seeks to increase the percentage of time that the principal spends on instruction, we explored whether there was a relationship between student achievement gains and either of the following:

- The *increase in the percent of time* the principal spent on instructional tasks, as defined and measured by the SAM project, over the first year of participation

- The *absolute percent of time* the principal spent on instructional tasks at the end of one year

We tested the correlation between increases in principal instructional time use and one-year and two-year student achievement gains. We found no cohorts of schools with correlations showing a statistically significant relationship between increases in principal time spent on instruction and one-year or two-year gains.

We also tested whether there was a correlation between the percent of time the principal spent on instruction at the end of one year and student achievement gains, with gains measured over one year and two years. We conducted the analyses separately for elementary and secondary schools, since the percent of time spent on instructional tasks was typically higher in elementary schools. In no cohort of schools in any of the states did a two-tailed test show statistical significance.[9]

In other words, among schools participating in the project, neither the principal's change in behavior nor the sheer amount of time spent on instruction after one year had a discernible relationship with school-level achievement gains as of that school year or the following year.

Finally, because the National SAM Project defines a gain of 13 percentage points in the principal's instructional time as successful implementation, we repeated all the comparative analyses presented in this report for the 25 schools in which the principals met that goal, in relation to their comparison schools. For the most part, these high-implementation schools did not fare significantly better in achievement than their matched comparisons (Exhibits 12 and 13). In one cohort these high-implementation schools had significantly higher one-year gains than their comparison schools: the four successful Iowa SAM schools had an average one-year increase of 2.2 percentage points in the proportion of students who were proficient, while proficiency in their comparison schools decreased by 4.5 percentage points. The difference was statistically significant (p=.006). Over two years, the successful Iowa SAM schools had an average two-year increase of 2.5 percentage points, compared with a decrease of 1.2 percentage points for comparison schools, but this difference was not statistically significant (p=.061). So, although four high-implementation Iowa SAM schools outperformed their matched schools after one year, the significant difference was not sustained over two years.

[9] Using a one-tailed test of significance, we found a difference in the Iowa schools: there was a positive correlation between principals' instructional time after one year in the project, and students' gains in reading at the second year after baseline testing. This result was not replicated in mathematics in that state or in any achievement measures in the other states. However, like the results reported above for the trends in these Iowa schools in relation to their comparison schools, it suggests that we cannot rule out the possibility of benefit to student achievement associated with implementation of the SAM project.

Exhibit 12: One-Year Change in Student Performance of High Implementation SAM Schools and Comparison Schools, By Cohort

Year One		High Implementation SAM Schools		Comparison Schools		Difference	
Cohort	**Measure**	**(n)**	**Average One-Year Change in Student Performance**	**(n)**	**Average One-Year Change in Student Performance**	**Mean Difference**	**p-value**
Kentucky Pilot	School Accountability Index Scores	(n=1)	▲ + 5.7	(n=1)	▲ + 6.5	▼ **- 0.8**	
Kentucky 2006 Base Year	School Accountability Index Scores	(n=3)	▲ + 4.3	(n=3)	▲ + 8.6	▼ **- 4.3**	(p=.262)
Kentucky 2007 Base Year	School Transition Index Scores	(n=15)	▲ + 2.2	(n=15)	▼ - 0.4	▲ **+ 2.6**	(p=.146)
Iowa 2007 Base Year [a]	Average Percent Proficient	(n=4)	▲ + 2.2	(n=8)	▼ - 4.5	▲ **+ 6.7**	(p=.006)
Illinois 2007 Base Year	PSAE Scores	(n=2)	▼ - 1.2	(n=3)	▼ - 3.0	▲ **+ 1.8**	(p=.585)

[a] *Cohort's one-year increases in achievement exceeded those of their matched comparison schools.*

Exhibit Reads: In the Kentucky pilot SAM school in which the principal's instructional time increased by at least 13 percentage points in the first year, there was a 5.7 point gain in student achievement over one year; this was 0.8 points less than the gain in that school's matched comparison. For other groups of schools, the statistical significance of the achievement comparison is shown in the exhibit. Note that one-year gains of the Kentucky 2006 base-year cohort should be interpreted with caution, due to changes in testing.

Exhibit 6: Two-Year Change in Student Performance of High Implementation SAM Schools and Comparison Schools, By Cohort

Year Two		High Implementation SAM Schools		Comparison Schools		Difference	
Cohort	**Measure**	**(n)**	**Average Two-Year Change in Student Performance**	**(n)**	**Average Two-Year Change in Student Performance**	**Mean Difference**	**p-value**
Kentucky Pilot	School Accountability Index Scores	(n=1)	▲ + 2.1	(n=1)	▼ - 2.0	▲ **+ 5.9**	
Kentucky 2006 Base Year	School Accountability Index Scores	(n=3)	▼ - 1.5	(n=3)	▲ + 8.6	▼ **- 10.0**	(p=.163)
Kentucky 2007 Base Year	School Transition Index Scores	(n=15)	▲ + 0.7	(n=15)	▼ - 0.1	▲ **+ 0.8**	(p=.746)
Iowa 2007 Base Year	Average Percent Proficient	(n=4)	▲ + 2.5	(n=8)	▼ - 1.2	▲ **+ 2.5**	(p=.061)
Illinois 2007 Base Year	PSAE Scores	(n=2)	▲ - 2.8	(n=4)	▼ - 2.6	▼ **- 0.2**	(p=.928)

Exhibit Reads: In the Kentucky pilot SAM school in which the principal's instructional time increased by at least 13 percentage points in the first year, there was a 2.1 point gain in student achievement over two years; this was 5.9 points more than the change in that school's matched comparison. For other groups of schools, the statistical significance of the achievement comparison is shown in the exhibit. Note that gains of the Kentucky 2006 and 2007 base-year cohorts should be interpreted with caution, due to changes in testing.

Conclusions

These analyses show a mixed picture on student achievement. While all cohorts of participating schools experienced a one-year increase in achievement following the introduction of the SAM project, they did not perform significantly better than their comparison schools, on average, over that first year. The two-year achievement gains of the Iowa and Illinois 2007 cohorts of SAM schools were significantly higher than their matched comparison groups. The same was not true, however, for other cohorts of SAM schools.[10]

Looking exclusively at those SAM schools judged successful because their principals registered a 13-percentage-point increase in instructional time, the four schools in one cohort performed significantly better than their comparison schools over one year, but the advantage was not sustained over two years, and none of the other cohorts (comprising 21 successful schools in total) had one-year or two-year gains significantly higher than their comparison schools.

A more conclusive assessment of project impact would require data from additional, larger cohorts of schools and more fine-grained examination of student-level data. Based on the available data, however, it is not clear that principals can bolster student gains by joining the SAM project or by increasing the percent of time they spend on the instruction-related tasks that the SAM project has defined.

[10] The two-year trends for Kentucky schools that joined the project in 2006 and in 2007 should be interpreted with caution, due to changes in testing. See note 6 above.