School Improvement Grant

Evaluation Report for 2010-2011

West Virginia Department of EDUCATION
Office of Research
Division of Curriculum and Instruction

School Improvement Grant

Evaluation Report for 2010-2011 (Redacted Edition)

Anduamlak Meharie

Nate Hixson

Patricia Cahape Hammer

West Virginia Department of EDUCATION
Office of Research
Division of Curriculum and Instruction

West Virginia Department of Education
Division of Curriculum and Instructional Services
Office of Research
Building 6-Room 722 State Capitol Complex
1900 Kanawha Boulevard East
Charleston, WV 25305
http://wvde.state.wv.us/

January 2012
(Redacted edition published in March 2013)

Jorea M. Marple
State Superintendent of Schools
West Virginia Department of Education

Robert Hull
Associate Superintendent
West Virginia Department of Education

Larry J. White
Executive Director
Office of Research

Prepared for
Jan Stanley
Executive Director
Office of Title I

Content Contact
Anduamlak Meharie, Ph.D.
Coordinator
Office of Research
ameharie@access.k12.wv.us

Suggested Citation

Meharie, A., Hixson, N., Hammer, P.C. (2012/2013). *School Improvement Grant: Evaluation report for 2010–2011 (redacted edition)*. Charleston, WV: West Virginia Department of Education, Division of Curriculum and Instructional Services, Office of Research.

Note about this edition: A portion of the data used in this study was collected with the understanding that it would be held confidential; consequently we have used codes in place of school names throughout this publically available version of the report (originally published for internal use only in 2012) and redacted information that could be identified with any individual school.

Executive Summary

This report provides the results of a multimethod evaluation research study to determine the impact of the School Improvement Grant program on participating schools in West Virginia. School Improvement Grants (SIGs) are awarded to state education agencies by the U.S. Department of Education under Section 1003(g) of the Elementary and Secondary Education Act of 1965, which was reauthorized as the No Child Left Behind Act in 2002. States provide subgrants to help local education agencies (LEAs) increase student achievement in Title I schools identified for improvement, corrective action, or restructuring. To be eligible for a SIG, an LEA's proposed improvement activities must be focused on the persistently lowest achieving school(s) located in that district.

In West Virginia, an application process led to eight districts, representing 15 schools, being approved for funding in August, 2010. As outlined in the West Virginia SIG application, the West Virginia Department of Education (WVDE) Office of Title I SIG program staff are working to achieve four broad project goals as they support LEAs in implementing the transformation model in their SIG schools:

Goal 1. Build the capacity of LEAs that receive SIG funding to successfully lead school improvement efforts in persistently low performing schools.

Goal 2. Strengthen teacher and leader effectiveness in SIG schools in order to improve the quality of instruction provided in schools.

Goal 3. Improve the academic achievement of students in SIG schools. All activities associated with the other goals support Goal 3.

Goal 4. Develop comprehensive systems of support in each SIG school and provide the structural capacity that is necessary for continuous school improvement.

This study was conducted by the WVDE Office of Research, in collaboration with the Office of Title I, to measure the extent to which the SIG program met the four broad project goals outlined above. Evaluators developed four broad evaluation questions (EQs) aligned to these project goals:

EQ1. To what extent has the SEA provided adequate support to LEAs and participating schools to successfully lead school improvement efforts?

EQ2. To what extent has the SIG program resulted in improved quality and quantity of instruction and nonacademic supports in participating schools?

EQ3. To what extent has the SIG program resulted in improved academic achievement among students in participating schools?

EQ4. To what extent have SIG schools successfully implemented and institutionalized practices and structures that are supportive of continuous school improvement?

Methodology

For EQ1, WVDE researchers reviewed documents related to professional development sessions and technical assistance events for the SIG program, and conducted an End-of-Year Survey of Title I directors working in SIG schools. The End-of-Year Survey was completed by eight representatives from seven districts near the end of the 2010-2011 year; it elicited feedback regarding the quality, relevance, and usefulness of professional development and technical assistance that districts received from the WVDE to help implement the Title I SIG program during the school year. The survey was also used to gather data on professional development and technical assistance that districts provided to schools implementing the SIG program during the same school year. Descriptive statistics were interpreted for survey responses.

For EQ2, researchers analyzed Instructional Practices Inventory (IPI) data collected onsite by Office of Title I program staff during the 2010-2011 school year. Fourteen of the original 15 Cohort 1 SIG schools participated. One school, Mount Hope High School, closed in June 2011, and did not take part in the IPI. Pearson's correlation coefficient (r) was used to measure the strength of relationships between the IPI data and proficiency rates in the four content areas (i.e., mathematics, reading/language arts, science, and social studies) tested by the West Virginia Educational Standards Test 2 (WESTEST 2) for SIG schools. We also compared baseline IPI data from SIG schools to *typical* baseline IPI data (i.e., data collected by IPI developers from schools across the nation). Additionally, we also analyzed results of the Effective School Practice Survey, conducted in May 2011, which was completed by WVDE Title I coordinators, LEA Title I directors, and school improvement specialists. Together they rated 11 of the 14 schools. Descriptive statistics were interpreted for the Effective School Practice Survey at the program level (aggregated across all SIG schools) and at the individual school level. Finally, we used Pearson's correlation coefficient (r) to measure the strength of relationship between Effective School Practice Survey data and proficiency rates in the four content areas tested by WESTEST 2 for SIG schools.

For EQ3, researchers collected available testing records for all students who were enrolled in the SIG and a matched set of non-SIG comparison schools during SY2009, SY2010, and SY2011 to analyze *program level effects*. For the main analysis, we used SY2010 data as the baseline point, given that it was the most recent testing point prior to SIG implementation. Once t tests indicated comparison school matching was successful, we conducted RM ANOVAs using panel data to detect the presence of any statistically significant interaction effects between time and group membership (SIG or non-SIG comparison) on mathematics and reading/language arts achievement. Our models included individual students who progressed up the grade span across 2 years of time (e.g., Grades 3–4, Grades 4–5, etc.). Follow-up descriptive analyses examined which group of students (i.e., SIG or non-SIG comparison) exhibited higher performance over time. For *school level effects*, we developed a descriptive school-level achievement profile for each Cohort 1 SIG school. These profiles examined the 3-year trend in overall school-wide proficiency rates from SY2009 to SY2011. Each profile contained the trend for the SIG school and the corresponding non-SIG comparison school selected for this study.

For EQ4, WVDE researchers analyzed survey data collected using the Measure of School Capacity for Improvement (MSCI), and data collected onsite by Office of Title I program staff using the School Culture Survey during the 2009-2010 and 2010-2011 school years. For both surveys, Pearson's correlation coefficient (*r*) was used to measure the strength of relationships between survey data and proficiency rates on the four content areas tested by the WESTEST 2 for SIG schools. Descriptive statistics were also interpreted for each of the domains of both surveys at the program level (aggregated across all SIG schools) and at the individual school level. All available professional staff from fourteen of the original 15 Cohort 1 SIG schools participated in the MSCI (n = 571) and School Culture Surveys (n = 396). The school that closed in June 2011 (Mount Hope High School), is omitted from this report.

Findings

As a result of these investigations, WVDE researchers made the following findings.

EQ1 Findings

Four major WVDE-sponsored professional development sessions (School Leadership Team Conferences) were provided during the 2010–2011 school year, which were attended by a total of 425 participants, including representatives from 11 Tier I state schools. Session themes included roles and responsibilities of school improvement specialists, school culture and data utilization, collaborative teams' curriculum (CSOs) and student leadership. Additionally, there were a total of 358 technical assistance reports posted on the WVDE Statewide System of Support (SSOS) websites for the 14 SIG schools. The majority of technical assistance provided to SIG schools focused on fostering collaboration, strategic planning, and the collection and utilization of various types of data to drive school improvement efforts.

Overall, respondents had favorable responses regarding the quality, usefulness, and relevance of the four *professional development sessions* and training materials provided during the sessions. Presenters were also well received in regards to their knowledge of the subject matter, clarity, effectiveness, and their helpfulness in answering questions during the sessions. Respondents indicated that *Statewide System of Support (SSOS) Teleconferences* were well organized, relevant, served as avenues of beneficial feedback, and helped improve communication among the different levels of their school system. Most respondents, however, regarded teleconferences as too time consuming and a couple of respondents commented on the fact that topics were more aligned with the interest of the state department instead of participants' own work in schools. *Technical assistance* provided by the Office of Title I to SIG schools was very highly received by respondents; and technical assistance providers were perceived to be very knowledgeable, committed, accessible, and instrumental to the school improvement process, particularly because technical assistance providers can attend to the needs of each SIG school directly. Survey responses also indicate that SIG schools have benefited from *technical assistance provided by districts*. Similar to technical assistance provided by the Office of Title I, comments indicate that this is particularly the case because services provided by districts can be individualized to schools in their districts. *Challenges* to school improvement efforts identified by respondents include issues around

the amount of time that some staff are required to be away from schools, defining the roles of various stakeholders, and the unwillingness of some staff to change.

EQ2 Findings.

Analysis of IPI data indicates that *teacher-led instruction* was found to be positively correlated with reading/language arts proficiency rates. *Student work with teacher engaged* was negatively correlated with proficiency rates for reading/language arts and science while *student work with teacher not engaged* was negatively correlated with proficiency rates for mathematics. The broad category of *disengagement* was found to be negatively correlated with mathematics proficiency rates. The combination of *student work with teacher not engaged* and *student work with teacher engaged* was negatively correlated with proficiency rates for reading/language arts and science. The grouping of the IPI categories *teacher-led instruction*, *student learning conversations*, and *student active engaged learning* was positively correlated with mathematics, reading/language arts, and science proficiency rates.

Overall, compared to the *typical* IPI data, SIG schools, at each of the three programmatic levels, showed similar or slightly higher percentages of less desirable IPI categories and, conversely, similar or slightly lower percentages of desirable IPI categories. The comparison at the group level, however, conceals considerable variation among individual SIG schools, a few of which have higher percentages of desirable categories and similar or lower percentages of less desirable categories. Compared to historically unsuccessful schools, SIG middle schools showed slightly higher percentages of desirable categories and similar or lower percentages of less desirable categories.

Results of the *Effective School Practice Survey* scores (for 11 of the 14 SIG schools) indicate that none of the mean domain scores reached 3.0 on the 4-point scale, which would have indicated strong fidelity to practices research has shown to be present in effective schools. This indicates that none of the domains are areas of particular strength for SIG schools as a whole at this point of implementation. The survey identified three areas of particular concern for SIG schools as a group: *parents and learning*; *curriculum, assessment, and instructional planning*; and *classroom instruction*. In these domains, 30% of schools or less had scores that met the threshold for above-average fidelity. According to analysis of Pearson's *r*, these three domains were also the domains that showed very strong positive relationships with proficiency rates and tested statistically significant at least at the 0.05 level.

EQ3 Findings

Program level results illustrated that, as far as the first SIG implementation year is concerned, there were relatively few significant differences between students in Cohort 1 SIG and non-SIG schools. However, some promising preliminary findings did emerge. We found that the students in Cohort 1 SIG schools outperformed their counterparts in non-SIG comparison schools by a statistically significant margin in the Grade 3–4 cohort in both reading/language arts and mathematics, and also in the Grade 5–6 cohort in reading/language arts. On the other hand, students in the non-SIG comparison schools outperformed their counterparts in Cohort 1 SIG schools by a statistically significant margin in the Grade 6–7 cohort in reading/language arts.

School level results indicated that both the SIG and non-SIG comparison schools have illustrated some extraordinary gains during the past 3 years (i.e., SY2009–SY2011). However, in both reading/language arts and mathematics, more SIG schools increased their proficiency rates over time in both reading/language arts and mathematics than non-SIG comparison schools. Furthermore, as a group, the Cohort 1 SIG schools had a median gain in reading/language arts that was 6.12% higher than the median gain for non-SIG comparison schools. For the same time period in mathematics, the SIG schools had a median gain that was 3.67% higher than the median gain for non-SIG comparison schools. Gains have also occurred for both groups during the span between SY2010 and SY2011 which represents the first year of SIG implementation. Our results illustrate that in reading/language arts, the same number of schools in both groups (SIG and non-SIG comparison) improved their proficiency rates during this time. Furthermore, the median reading/language arts proficiency gain for Cohort 1 SIG schools during this period was very similar to non-SIG comparison schools (a difference of only 0.4%). However, with respect to mathematics, we found that approximately 78% of SIG schools increased their proficiency rates over the SY2010–SY2011 period while only 50% of non-SIG comparison schools were able to do so. The Cohort 1 SIG schools had a median gain of 5.4% in mathematics while the non-SIG comparison schools exhibited a median gain that was less than 1% (a difference of 5.3%).

EQ4 Findings

Analysis of *MSCI survey* data identified at least two critical areas of need for SIG schools as a group. These are *capacity for peer reviewed practice* and *increasing expectations for student performance*. These were areas where the median for SIG schools fell very near the 50th percentile. The MSCI identified three areas of strength for SIG schools as a group. These are *differentiated instruction, technical resources,* and *improvement program coherence*. These are areas where the median for SIG schools fell near or above the 70th percentile. There were, however, a few schools that rated quite low in these domains, and their performance will continue to need monitoring and intervention.

Results of the *School Culture Survey* identified at least three areas of strength for SIG schools as a group: *unity of purpose*, *collegial support*, and *professional development*. These are areas where domain average scores for SIG schools fell above the overall mean for all seven dimensions. There were, however, a few schools that rated quite low in these domains, and their performance will continue to require monitoring and intervention. The survey also identified at least three areas of need for SIG schools as a group: *learning partnerships*, *teacher collaboration*, and *collaborative leadership*. These domains had average scores for SIG schools that fell very near or below the overall mean for all seven dimensions. Two of these domains, *teacher collaboration* and *collaborative leadership*, were also the most strongly correlated with student achievement data on WESTEST 2.

Recommendations

The WVDE Office of Research makes the following broad recommendations for implementation, ongoing monitoring, and evaluation of the SIG program based on findings presented in this report:

Implementation

Based on document reviews and responses to the various surveys conducted for this evaluation, there were some areas of practice that show movement in a positive direction, and we therefore encourage the Office of Title I to continue to—

- Link professional development and technical assistance to each other and to cohesive and well-defined goals for improvement, as was done during the 2010–2011 school year.
- Provide professional development that LEA Title I directors perceive as high quality, useful, and relevant, and as producing positive change in their own knowledge, behaviors, and beliefs and attitudes.
- Maintain the quality of support provided by districts in 2010-2011. Evidence from this study suggests that, when compared to professional development, district-provided support affords greater opportunity to differentiate intervention activities based on specific school needs.

Other evidence from this study suggests where efforts need to be intensified, continued, or adjusted:

- Use SSOS teleconferences, as much as possible, as two-way conversations to elicit information about participants' *current work* and to provide feedback that would help guide their school improvement efforts.
- Use more onsite technical assistance, as well as district level services, to help alleviate one of the most recurring concerns expressed by SIG school staff—that is, time spent away from their school buildings participating SIG activities.
- Provide additional guidance to elucidate the roles and responsibilities of all those involved in the school improvement effort in SIG schools and their home districts.
- Provide guidance on appropriate ways to handle staff members who are resistant to efforts to bring about positive changes in SIG schools.
- Encourage all SIG schools to collect IPI data multiple times per year.
- Provide guidance on ways schools can increase student involvement in *higher order/ deeper learning activities* (IPI Categories 5 and 6), and reduce the incidence of *student work with teacher not engaged* (IPI Category 2) and *student work with teacher engaged* (IPI Category 3).
- Focus additional support on addressing effective school practices, which evidence from the Effective School Practice Survey suggests need more attention. Domains that seem to need the most attention by the greatest number of schools include *parents and learning*; *curriculum, assessment, and instructional planning*; and

classroom instruction. These domains not only had the lowest mean scores in SIG schools, but were also the most strongly correlated with student achievement data on WESTEST 2, suggesting that improvement in these areas may have a tangible impact on student achievement.

- Focus additional support to address deficiencies in SIG schools related, especially, to the MSCI subscales *peer reviewed practice* (i.e., the observation and review by staff of their peers' work), and *expectations for student performance* (i.e., staff members' expectations of the students and their beliefs that all students can perform well academically).
- Focus additional support on addressing widespread deficiencies in SIG schools in the School Culture Survey domains of *learning partnerships* (i.e., the degree to which teachers, parents and students work together for the common good of the student), *teacher collaboration* (i.e., the degree to which teachers engage in constructive dialogue that furthers the educational vision of the school), and *collaborative leadership* (i.e., the degree to which school leaders establish and maintain collaborative relationships with school staff). Two of these domains, *teacher collaboration* and *collaborative leadership*, not only had the lowest mean scores, but were also the most strongly correlated with student achievement data on WESTEST 2. Improvement in these areas could potentially result in better student achievement.

Ongoing Monitoring and Evaluation

- Reserve holistic judgments of the SIG program's impact upon student achievement until achievement data spanning all years of the SIG program are collected and analyzed thoroughly. Our interim evaluation of student achievement in SIG schools should be used diagnostically as a "temperature check" of the SIG program and to identify schools where relatively minor gains are occurring when compared with other SIG schools and matched non-SIG schools. Using these data in this manner allows WVDE to intervene in these sites before an opportunity for course correction—using the substantial resources provided by the SIG—is lost.
- Identify a more robust matching methodology for our final round of program-level analyses, one that allows for adequate student-level matching at all programmatic levels. Accomplishing this degree of a match will be absolutely critical if we wish to provide a fair and accurate view of the SIG program. The Office of Research in consultation with Title I staff must also consider whether a longitudinal panel data approach is the most appropriate method to evaluate the SIG program given the level of attrition that occurs with student records.
- Investigate whether or not the program level effects on WESTEST 2 mathematics and reading/language arts scores found in elementary schools reflect some set of innovative practices. If so, these practices should be identified, nurtured, and, if possible, replicated in middle/high schools.
- Continue to encourage SIG schools to administer the School Culture Survey on an annual basis and utilize the data for school improvement efforts.

- Utilize average percentages of IPI data collected at least two times during each school year in future evaluations.
- In addition to LEA Title I coordinators, extend the invitation to participate in the 2011-2012 End-of-Year Survey to school leadership team members and school improvement/transformation specialists, who can contribute valuable feedback regarding the quality, relevance, and usefulness of professional development and technical assistance schools and districts receive during the school year.
- Encourage all school improvement specialists to participate in the Effective School Practice Survey. It is clear from the frequency of "I don't know" responses from SEA Title I directors that school improvement specialists are best situated to provide relevant feedback on many of the survey items included in the survey.
- Re-examine the correlation between the Effective School Practice Survey and WEST-EST 2 proficiency once this wider audience has participated.
- Eliminate MSCI survey from analysis at least until the final year of SIG program for the following reasons: (a) there is a tremendous amount of overlap among the MSCI, School Culture, and Effective School Practice surveys; (b) the MSCI survey is best used as a diagnostic instrument at the onset of program implementation to gauge schools' capacity for improvement and perhaps at the end of a program cycle to measure any gains obtained as a result of improvement efforts; and (c) eliminating the MSCI survey will reduce the burden of completing an additional survey by school staff in SIG schools who are already tasked with a great many responsibilities.

Limitations

This study has a number of limitations that the reader should be aware of: (a) the study relies heavily on perceptual survey data, therefore, analysis and interpretations are only inclusive of the views expressed by those individuals who were surveyed—furthermore, in most cases we did not reach the entire population of the target group(s); (b) in most cases, the sample size of participants (e.g., number of SIG schools, End-of-Year survey respondents) was too small for methods of inferential statistics, rendering the data primarily descriptive in nature; (c) results and recommendations in this study are based on only 1 year of implementation and data collection; (d) technical issues including the failure to match adequately those students in SIG and non-SIG middle schools, the decision to use longitudinal panel data for program-level analyses, and low observed statistical power for some analyses limit our confidence in the results of the interim student achievement analyses; (e) our evaluation design was not adequate to provide very detailed information about the various implementations of the SIG program as they vary substantially across the 14 Cohort I schools; and (f) school-level achievement comparisons among SIG and non-SIG schools do not represent rigorous statistical comparisons and school-wide proficiency rates are not analogous to those reported for accountability purposes. Due to these limitations, the reader should use some caution when considering the results of this study. For a detailed look at limitations of the study, refer to the discussion section starting on page 101.

Contents

List of Figures

List of Tables

Introduction

School Improvement Grants (SIGs) are awarded to state education agencies by the U.S. Department of Education under Section 1003(g) of the Elementary and Secondary Education Act of 1965, which was reauthorized by the No Child Left Behind Act in 2002. States provide subgrants to help local educational agencies (LEAs) to improve student achievement in Title I schools identified for improvement, corrective action, or restructuring. To be eligible for a SIG, an LEA's proposed improvement activities must be focused on the persistently lowest achieving school(s) located in that district (U.S. Department of Education, 2010a).

The School Improvement Grant (SIG) Program

To identify the persistently lowest achieving schools, the West Virginia Department of Education (WVDE) began by computing for every public school in the state, the percentage of all students[1] scoring at the *proficient* level on WESTEST 2 reading/language arts and mathematics assessments in Grades 3 through 8 and 11.[2] The schools were then rank ordered from lowest to highest achieving. Equal weight was also given to schools' lack of progress—defined as two consecutive years of not making adequate yearly progress (AYP) in the *all student* group for school years 2007–2008 and 2008–2009 (WVDE, n.d.).

The U.S. Department of Education (USED) classifies low achieving schools as falling into three categories, or *tiers*. West Virginia, in keeping with the USED classification system, categorized the state's persistently lowest achieving five Title I schools into Tier I or lowest achieving 5%.[3] Notably, West Virginia has no high schools in the state with persistent graduation rates below 60%, as described in federal regulations (34 C.F.R. §200.19(b)); consequently, there were no high schools added to the Tier I list.

In identifying the persistently lowest achieving schools to be categorized into Tier II, West Virginia selected the secondary schools that are eligible for, but do not receive Title I funds, and that are among the lowest achieving 5% of secondary schools. All schools identified for Tier II must also have a percentage of poverty that is both above the respective district poverty rate and above 40% (USED, 2010b). Notably, none of the schools added to the newly eligible list exhibited higher achievement than the highest achieving Tier II schools as identified by the SEA under the December 10th, 2010 SIG final requirements.

The schools on the Tier III list included the remaining 18 Title I schools identified for improvement, corrective action, or restructuring as required by guidance from USED

[1] Scores for students with limited English proficiency and students with disabilities, including those who take an alternative assessment based on alternate academic achievement standards, were included in computing the percentage proficient values for the all student group.

[2] West Virginia has no public charter schools.

[3] Tier I encompassed the bottom five schools in the aforementioned rank-ordered list.

(2010b). The option to add schools to Tier III was not exercised in order to adequately fund programs of sufficient size and scope in Tier I and Tier II schools.

LEA grant application process

In 2009, 20 of West Virginia's 57 school districts (35%) had schools that met the aforementioned criteria to be classified as persistently lowest performing schools. Of these, 17 districts elected to apply for a SIG by the application closing date of June 14, 2010. These applications represented 27 schools classified as Tier I, II, or III. Each LEA application was judged on the quality of the proposed plan, the level of need for funding—with priority given to districts with Tier I and Tier II schools—and the capacity of the district to make substantive changes to improve student achievement. Finalists gave oral presentations to the grant review committee before the final awardees were chosen. Feedback from the presentations and the grant reviewers was used by the districts to revise their final grant applications. After a second review, eight districts representing 15 schools were approved for funding in August, 2010 (See Table 1).

Table 1. Districts and Schools Selected for School Improvement Grants

District	School	School NCES Identification Number	Tier
Berkeley	Martinsburg North Middle School	540006000024	II
Doddridge	Doddridge County Elementary School	540027001059	I
Fayette	Mount Hope High School	540030000195	II
Kanawha	Malden Elementary School	540060000416	I
	Cedar Grove Middle School	540060001252	II
	East Bank Middle School	540060000386	II
	Riverside High School	540060001043	II
	Stonewall Jackson Middle School	540060001442	II
Lincoln	West Hamlin Elementary School	540066000502	I
	Hamlin PK-8 School	540066001237	II
	Guyan Valley Middle School	540066001242	II
McDowell	Mount View High School	540081001246	II
	Sandy River Middle School	540081001046	II
Roane	Geary Elementary/Middle School	540132001305	II
Wood	Franklin Elementary Center	540162000679	I

Source. West Virginia Department of Education (2010d). West Virginia funded districts summary: School Improvement Grant 2010-2011, Tier I and Tier II schools, Cohort 1. Charleston, WV: Author.

Program implementation

Implementation began in the schools in August, 2010 and will extend through the 2012–2013 school year. According to federal SIG program requirements, LEA applicants are required to implement one of four school improvement models: turnaround, restart, school

closure, or transformation (USED, 2010a). West Virginia applicants are all implementing the transformation model.[4] USED guidance describes this model as follows:

> An LEA implementing a transformation model must
>
> 1. Replace the principal who led the school prior to commencement of the transformation model;
> 2. Use rigorous, transparent, and equitable evaluation systems for teachers and principals that
>   a. Take into account data on student growth as a significant factor as well as other factors, such as multiple observation-based assessments of performance and ongoing collections of professional practice reflective of student achievement and increased high school graduation rates; and
>   b. Are designed and developed with teacher and principal involvement;
> 3. Identify and reward school leaders, teachers, and other staff who, in implementing this model, have increased student achievement and high school graduation rates and identify and remove those who, after ample opportunities have been provided for them to improve their professional practice, have not done so;
> 4. Provide staff ongoing, high-quality, job-embedded professional development that is aligned with the school's comprehensive instructional program and designed with school staff to ensure they are equipped to facilitate effective teaching and learning and have the capacity to successfully implement school reform strategies; and
> 5. Implement such strategies as financial incentives, increased opportunities for promotion and career growth, and more flexible work conditions that are designed to recruit, place, and retain staff with the skills necessary to meet the needs of the students in a transformation model. (USED, 2010a, pp. 26-27)

While proposed activities vary among the LEA SIG awardees, all plans remain within the guidelines for the transformation model.

As outlined in the West Virginia SIG application, the Office of Title I SIG program staff are working to achieve four broad project goals as they support LEAs in implementing the transformation model in their SIG schools:

[4] Two of the models, the school closure and turnaround options, are precluded by the rural nature of the state. More than half of schools and approximately 40% of students are in rural areas. Additionally 25 of the 55 county districts have only one high school, making the closure of a high school in those districts untenable. It would also be difficult to replace principals and large numbers of teachers in rural schools and districts, which already struggle to fill positions with highly qualified educators. The restart model involves closing a school and reopening it as a charter school. West Virginia law does not provide for charter schools (WVDE, 2010b). These limitations left the transformation model as the only viable option.

Goal 1. Build the capacity of LEAs that receive SIG funding to successfully lead school improvement efforts in persistently low performing schools through the following activities:

a. Engage external supporting partners to help LEAs build their capacity to support the transformation of struggling schools.

b. Realign and expand the current capacity of the state system of support structure to monitor the process of transformation at all struggling schools in the state and build capacity at the Regional Education Service Agency (RESA) and LEA levels.

c. Build strong teams to support struggling schools at the RESA, LEA, and school levels.

d. Design a Whole Child Early Warning System and a Whole School Early Warning System that enable multiple users to use data to drive the school improvement process.

e. Utilize an evaluator to design a rigorous evaluation and report formative results annually for 2 years and summative results at the end of year 3. (WVDE, 2010e, p. 33)

Goal 2. Strengthen teacher and leader effectiveness in SIG schools in order to improve the quality of instruction provided in schools through the following activities:

a. Utilize criteria developed by the West Virginia State System of Support to assist low performing schools in the selection of a school-based transformation specialist.

b. Implement the structures, supports, and professional development that teachers need to be successful in professional learning communities.

c. Provide professional development designed to assist teachers in implementing the Content Standards and Objectives utilizing standards-based instructional strategies. (WVDE, 2010e, p. 33)

Goal 3. Improve the academic achievement of students in SIG schools. All activities associated with the other goals support Goal 3.

Goal 4. Develop comprehensive systems of support in each SIG school and provide the structural capacity that is necessary for continuous school improvement through the following activities:

a. Provide a school-wide system of differentiated supports for struggling students and students with disabilities.

b. Establish school-based case management teams to identify non-academic issues for struggling students and then align the appropriate supports and services to the students' needs. (WVDE, 2010e, p. 33)

Purpose of the Study

The WVDE Office of Title I is responsible for fulfilling USED annual reporting requirements via the Consolidated State Performance Report (CSPR), which will cover all schools that receive SIG funds.

In this separate, internal evaluation of the SIG program, the Office of Research, in collaboration with the Office of Title I, will measure the extent to which the SIG program meets the four broad project goals outlined above. Evaluators have developed four broad evaluation questions (EQs) aligned to these project goals:

EQ1. To what extent has the SEA provided adequate support to LEAs and participating schools to successfully lead school improvement efforts?

EQ2. To what extent has the SIG program resulted in improved quality and quantity of instruction and nonacademic supports in participating schools?

EQ3. To what extent has the SIG program resulted in improved academic achievement among students in participating schools?

EQ4. To what extent have SIG schools successfully implemented and institutionalized practices and structures that are supportive of continuous school improvement?

Methods

Evaluation Question 1

EQ1. To what extent has the SEA provided adequate support to LEAs and participating schools to successfully lead school improvement efforts.

To investigate this question, we reviewed documents associated with the SIG program and conducted an End-of-Year Survey of Title I directors working in SIG schools. For the document review, we engaged in the following data collection and analysis activities:

- *Professional development sessions*. We requested professional development agendas and sign-in sheets for the 2010-2011 school year from the Office of Title I. Additionally, we consulted a number of Title I program personnel for additional data and for clarifications. Professional development agendas provided data on the date, location, and theme of each session. Sign-in sheets were reviewed to determine the number and roles/titles of session participants.
- *Technical assistance events*. We reviewed technical assistance reports associated with all 14 SIG schools in the 2010-2011 school year. The reports were posted on the WVDE Statewide School System of Support (SSOS) websites for each SIG school. The reports provided data on the names of technical assistance providers, date and location of the technical assistance events, summary of the assistance provided, and the intended audience for the assistance.

The methodological approach used for the End of Year Survey was as follows.

Participant characteristics

The End-of-Year Survey was completed by representatives from seven districts (Wood County did not participate in the survey). Two representatives from Doddridge County, a Title I director and a transformation specialist, participated in the survey. Therefore, a total of eight respondents completed the survey. Three out of the eight respondents, or 37.5%, indicated that they had occupied their current positions for less than a year and an additional 37.5% (n = 3) said they had been in their current positions between 1 to 5 years (Table 2). The remaining two respondents (25%) had been in their current position for at least 6 years. All respondents (100%) had at least 6 years of experience in education, with 87.5% (n = 7) having more than 15 years.

Most of the respondents had attended the conferences, the quality of which, they were being asked to evaluate. All four of the major WVDE-sponsored professional development sessions provided during school year 2010–2011 were attended by at least six of the eight respondents (75%). Three of the four professional development sessions were attended by at least seven of the respondents (87.5%), while one of the sessions (October 18-29, 2010) was attended by all eight respondents (100%).

Table 2. Characteristics of End-of-Year Survey Respondents

Respondent background	Response	Frequency	Percent
County	Berkeley	1	12.5
	Doddridge	2	25.0
	Fayette	1	12.5
	Kanawha	1	12.5
	Lincoln	1	12.5
	McDowell	1	12.5
	Roane	1	12.5
Position title	Transformation specialist	1	12.5
	Title I director	7	87.5
Number of schools in district implementing SIG	1	3	37.5
	2	2	25.0
	3	1	12.5
	5	1	12.5
	6	0	0.0
	More than 6	1	12.5
Number of years of experience in current position	Less than 1	3	37.5
	1 to 5	3	37.5
	6 to 10	1	12.5
	11 to 15	0	0.0
	More than 15	1	12.5
Number of years of experience in education	Less than 1	0	0.0
	1 to 5	0	0.0
	6 to 10	1	12.5
	11 to 15	0	0.0
	More than 15	7	87.5
WVDE-sponsored professional development sessions attended during school year 2010–2011?*	August 16–18, 2010 (Bridgeport Conference Center)	6	75.0
	October 18–20, 2010 (Bridgeport Conference Center)	8	100.0
	February 16–17, 2011 (Bridgeport Conference Center)	7	87.5
	April 7–8, 2011 (Embassy Suites in Charleston)	7	87.5

*Denotes items for which respondents could select all response categories that apply.

Sampling procedures

Researchers aimed for one district-level respondent from each of the eight districts.

Measures and covariates

We designed the Title I School Improvement Grant (SIG) End-of-Year Survey in consultation with staff the Office of School Improvement and the Office of Title I. It was intended to elicit feedback regarding the quality, relevance, and usefulness of professional development and technical assistance that districts received from the WVDE to help implement the Title I SIG program during school year 2010-2011, including the quality of the training materials and resources, the quality of the presenters, and a judgment about the impact of the professional development sessions. Additionally, the questionnaire asked

about the quality of the statewide system of support teleconferences; and the technical assistance they received, including the quality of the technical assistance providers. The survey was also used to gather data on professional development and technical assistance that districts provided to schools implementing the SIG program during the same school year; and the benefit to schools from services provided by districts. The survey was posted online using SurveyMonkey.

Research design

In mid May, 2011, an online survey invitation was sent via e-mail to Title I directors in the eight districts with SIG schools (Berkeley, Doddridge, Fayette, Kanawha, Lincoln, McDowell, Roane, and Wood). Data from the survey were tabulated and descriptive statistics were interpreted (see Results section, page 24).

Evaluation Question 2

EQ2. To what extent has the SIG program resulted in improved quality and quantity of instruction and nonacademic supports in participating schools?

To investigate this question, we analyzed data collected onsite by Office of Title I program staff, who used the Instructional Practices Inventory (Painter & Valentine, 1996) during the 2010-2011 school year. We also analyzed results of a survey, the Effective School Practice Survey, conducted in May 2011, which was based on items developed by the Center on Innovation & Improvement (Walberg, 2007) and the Academic Development Institute (Redding, 2006).

Participant characteristics

Fourteen of the original 15 Cohort 1 SIG schools participated in the Instructional Practices Inventory (IPI) component of the study. One school, Mount Hope High School, closed in June 2011, and did not take part in the IPI.

For the Effective School Practice Survey, five WVDE Title I coordinators, eight LEA Title I directors, and eight school improvement specialists working with SIG schools participated. Together they rated 11 of the 14 schools.

Sampling procedures

The full set of Cohort 1 schools and classrooms (for IPI observations), and all Title I directors (state and local) and school improvement specialists associated with SIG schools (for Effective School Practice Survey) were included in these two components of the study, so no sampling procedures were needed.

Measures and covariates

Instructional Practices Inventory

The IPI was developed by Bryan Painter and Jerry Valentine in 1996 to be used in Missouri for a multiyear comprehensive systemic school reform project called Achieving Success through School Improvement Site Teams (ASSIST). The process was later revised by Valentine in 2002, 2005, and 2007.

The IPI is a systematic process to profile school-wide student engaged learning on three broad IPI categories; *student engaged instruction*, *teacher-directed instruction*, and *disengagement*. The three broad categories are further broken down to six coding categories (Table 3) so that specific data about student engagement can be gathered and faculty can use it for instructional improvement. The six categories are nonhierarchical and they should simply be thought of as ways to categorize student engagement.

Table 3. Instructional Practices Inventory Categories

Broad Categories	Coding Categories
Student engaged instruction	(6) Student active engaged learning
	(5) Student learning conversations
Teacher-directed instruction	(4) Teacher-led instruction
	(3) Student work with teacher engaged
Disengagement	(2) Student work with teacher not engaged
	(1) Complete disengagement

The premise of the IPI process is that meaningful engagement (both student and teacher) in the learning/teaching process is vital for student academic success and is associated with increased student achievement. Based on analysis of IPI engagement data collected between 1996 and 2008, Collins and Valentine found, among other observations: (a) a significant relationship between the degree to which students are engaged in higher-order/deeper learning (IPI Categories 5 and 6) and student achievement scores on high-stakes accountability assessments; (b) a strong inverse relationship between disengagement (Category 1) and student achievement in the content areas of communication arts and mathematics; and (c) based on academic achievement, significant differences in types of student engagement between very successful and very unsuccessful middle schools (Valentine 2011).

Effective School Practice Survey

The Effective School Practice Survey was constructed with selections from a list of indicators for continuous school improvement found on the Center on Innovation & Improvement website (www.centerii.org). Evaluators from the Office of Research and WVDE Title I coordinators worked together to select a total of 48 indicators for five domains of Effective School Practice (see full list of indicators in Appendix B, Table B- 1, page 123). The number of indicators for each domain ranged between 3 and 15, as follows:

1. Leadership and decision-making (14)
2. Professional development (3)
3. Parents and learning (6)
4. Curriculum, assessment, and instructional planning (10)
5. Classroom instruction (15)

The group also worked on defining the verbal anchor values for a 4-point Likert-type response scale. Two additional categories, *I don't know* and *Too early to tell* were also included, but not used to calculate mean scores for the items. For the majority of items the 4-point Likert-type response scale was as follows: 1 = *Rarely, if ever*, 2 = *Sometimes,* 3 = *Often*, and 4 = *Consistently*. Nine survey items used 4-point Likert-type response scales that were item specific. A higher mean score (3.0 or above) indicates that the school is implementing effective school practice(s) with high fidelity (1 = *Little or no fidelity*, 2 = *Limited fidelity*, 3 = *Above average fidelity*, and 4 = *Strong fidelity*). The survey was then posted online (via SurveyMonkey) to facilitate a better response rate and ease the burden of data analysis.

Research design

Instructional Practices Inventory data collection and analysis

Trained observers collected IPI data through direct classroom observations. IPI protocols called for observers to use a map to systematically move throughout the school to ensure that data were collected proportionately from all instructional settings. Observers focused on the students' learning experiences during the first few moments of the observation (1-3 minutes) and coded it anonymously using the six categories of student engaged learning (see Table 3). Observers collected data continuously throughout the school day following the same systematic pattern so that each classroom was observed multiple times. The expected minimum number of observations per day is 100, although 125-150 observations per day are preferred and more typical.

Observations took place during a typical day (no unusual circumstances), avoiding Fridays when possible. Observers did not code classrooms during the first and last 5 minutes of a class period at the middle or high school level, or during content transitions at the elementary level. When coders observed a learning experience that fell between two categories, they recorded the category that represents the more favorable learning experience. Coders also designated *core* and *noncore* classes when they recorded their observation data. Later a *total* profile for all observations was generated by combining core and noncore observation data. Core classes included language arts, mathematics, science, and social studies. All other content areas were designated as noncore classes.

All IPI data for SIG schools used in this report came from the first data collection conducted by program staff from the Office of Title I during the 2010-2011 school year. Although in some SIG schools IPI data were collected more than once during the same school year through additional observations, the only set of IPI data available for *all* SIG schools for the 2010-2011 school year was the first one. Consequently, we decided to look only at data that were available for all schools and use it as a baseline for SIG schools.

For 11 of the 14 SIG schools the first IPI data were collected between September and December of 2010 (Table 4). Collection of IPI data for the other three SIG schools (all from the same county) was delayed as program staff waited to obtain permission from the county. IPI data for these three schools were collected between January and March, 2011.

Table 4. Time of First IPI Data Collections

School	1st IPI data collection
Cedar Grove Middle	September, 2010
Doddridge County Elementary	October, 2010
East Bank Middle	December, 2010
Franklin Elementary Center	October, 2010
Geary Elementary/Middle	September, 2010
Guyan Valley Middle	March, 2011
Hamlin PK-8	January, 2011
Malden Elementary	December, 2010
Martinsburg North Middle	October, 2010
Mount View High	October, 2010
Riverside High	October, 2010
Sandy River Middle	October, 2010
Stonewall Jackson Middle	September, 2010
West Hamlin Elementary	February, 2011

We used Pearson's correlation coefficient (r) to measure the strength of relationships between the IPI data and proficiency rates on the four content areas tested by the WESTEST 2 for SIG schools. WESTEST 2 is a custom-designed assessment for West Virginia students. The individual content assessments measure a student's levels of performance on clearly defined standards and objectives and skills. Student scores are based on test questions that have been developed and aligned to the West Virginia 21st Century CSOs.

We also compared baseline IPI data from SIG schools to what Collins and Valentine (Valentine, 2005) call *typical* baseline data—that is, IPI profiles from schools that have participated in projects undertaken by the Middle Level Leadership Center and from data voluntarily submitted to the Center by schools using the IPI process.

Findings from these analyses can be found in the Results section, beginning on page 34.

Effective School Practices Survey data collection and analysis

In mid-May 2011, a link to the online survey was sent to five WVDE Title I coordinators, eight LEA Title I directors, and eight school improvement specialists working with SIG schools. They were provided with a list of 21 SIG schools to rate using the survey (response rates from the three groups of respondents are summarized in Table 5). The 21 SIG schools were made up of 15 schools from Cohort 1 with an obligation date starting in October, 2010 and six schools from Cohort 2 with an obligation date starting in June, 2011 (Table 6). Although data from two Cohort 2 SIG schools were available they were nevertheless excluded from the analysis in this report. Their results will only be used as a baseline for future analysis. This report is based on available data from 11 Cohort 1 SIG schools. Schools not rated included Cedar Grove Middle, East Bank Middle, and Riverside High.

Table 5. Effective School Practice Survey School Ratings by Respondents' Position Title

	School ratings received	
Position Title	n	%
Total	25	100.0
SEA Title I coordinator	12	48.0
LEA Title I coordinator	9	36.0
School improvement specialist	4	16.0

Table 6. List of Schools Rated Using the Effective School Practice Survey and Number of Raters per School

School	County	Cohort	Rated	Number of raters per school
Cedar Grove Middle	Kanawha	1	No	NA
Doddridge County Elementary	Doddridge	1	Yes	1
East Bank Middle	Kanawha	1	No	NA
Franklin Elementary Center	Wood	1	Yes	2
Geary Elementary/Middle	Roane	1	Yes	3
Guyan Valley Middle	Lincoln	1	Yes	4
Hamlin PK-8	Lincoln	1	Yes	1
Malden Elementary	Kanawha	1	Yes	2
Martinsburg North Middle	Berkeley	1	Yes	2
Mount Hope High*	Fayette	1	Yes	1
Mount View High	McDowell	1	Yes	2
Richwood High	Nicholas	2	No	NA
Riverside High	Kanawha	1	No	NA
Romney Elementary	Hampshire	2	No	NA
Sandy River Middle	McDowell	1	Yes	2
Southside K-8*	McDowell	2	Yes	1
Spencer Elementary	Roane	2	No	NA
Stonewall Jackson Middle	Kanawha	1	Yes	1
Watts Elementary	Kanawha	2	No	NA
Welch Elementary*	McDowell	2	Yes	1
West Hamlin Elementary	Lincoln	1	Yes	1

*Eliminated from analysis: Mount Hope High closed school in June, 2011; Southside K-8 and Welch Elementary are Cohort 2 schools

Descriptive statistics were interpreted for each of the five domains at the program level (aggregated across all SIG schools) and at the individual school level. Additionally, we used Pearson's correlation coefficient (*r*) to measure the strength of relationship between Effective School Practice Survey data and proficiency rates on the four content areas tested by the WESTEST 2 for SIG schools.[5] Findings from all of these these analyses can be found in the Results section, beginning on page 44.

Evaluation Question 3

EQ3. To what extent has the SIG program resulted in improved academic achievement among students in participating schools?

[5] For help with interpreting Pearson's r, see the explanation provided in the Instructional Practices Inventory Results section (page 33).

To address EQ3, we conducted two analyses—one using student-level assessment data, which examined effects of the SIG program on academic achievement statewide; and another conducted at the school level, which provided information about each individual SIG school. Both analyses were conducted using the same sample of SIG and non-SIG comparison schools and WESTEST 2 scores for mathematics and reading/language arts.

Participant characteristics

Students whose scores were included in this analysis (SIG and non-SIG comparison schools) were primarily White (88%) or Black (10%), with a small population of Hispanic students (~1%). All other races/ethnicities represented less than 0.5%. Most students (65%) were eligible for free or reduced-price lunch and approximately 13% were eligible for special education services. Approximately 51% were male and 48% female. The grade-level breakdown for the sample of students appears in Table 7.

Table 7. Distribution of Program-Level Sample by SY2010 Grade

SY2010 Grade	*N*	% of Sample
Total	5,924	100
3	632	10.7
4	499	8.4
5	383	6.5
6	1,489	25.1
7	1,451	24.5
8	168	2.8
9	700	11.8
10	602	10.2

Sampling procedures

To select the set of schools to serve as a comparison group for Cohort 1 SIG schools, we considered eight matching criteria as described in Table 8. For all matches, we first attempted to meet the prescribed tolerances; however, we prioritized the match among mathematics and reading/language arts proficiency rates given that our primary intent in matching upon the demographic variables was to control for their potential influence upon student achievement.

Table 8. Criteria and Tolerances Used for Matching Schools

Matching Criterion	Tolerance Used
Programmatic level	Matched school had to be same programmatic level for NCLB accountability reporting (e.g., elementary, middle, secondary). Where possible, grade spans were matched between schools.
Enrollment	Enrollment had to match within 100 students between the two schools.
Free/reduced price lunch (FRPL) eligibility	10% tolerance for free and reduced-price lunch eligibility
Racial/ethnic diversity	10% tolerance for diversity
Special education eligibility	10% tolerance for special education eligibility
Mathematics proficiency (2010)	10% tolerance for mathematics proficiency rates
Reading/language arts proficiency (2010)	10% tolerance for reading/language arts proficiency rates
Title I eligibility	Matched school must be eligible for Title I

Using the tolerances described above, we were unable to select suitable matches for five of the Cohort 1 SIG schools (about 36%). We presume that this was due to the fact that these schools were among the neediest and lowest performing schools in the state, most of which were more racially/ethnically diverse than remaining schools in the state. As such, we

selected the closest available matches for these schools, always considering the match upon the proficiency variables to be the priority. We then consulted with the Office of Title I to verify that our final selections met their approval and to ensure that comparison schools, to their knowledge, did not differ from their matched counterparts on any substantively important, but unmeasured qualitative basis. We received no revisions to our list from Title I. The final set of matched schools appears in Table 9. SIG schools appear in italicized text and violations of the matching tolerances appear in bold red text.

Table 9. Summary of School Matching Results

School Name	Type	Grade Span	Enrollment*	FRPL %*[1]	Diversity %*	Sp Ed %*[2]	Reading/ Mathematics % 2010*[3]	Reading/ Language Arts % 2010*[3]
School I 1	M	6–8						
School I 2	M	6–8						
Group Δ			26	4.74	4.76	7.9	-1.93	1.9
School N1	E	PK–4						
School N2	E	PK–5						
Group Δ			9	**11.07**	-2.46	11.44	-3.07	2.08
School K1	E	PK–5						
School K2	E	PK–5						
Group Δ			-71	**-12.29**	5.87	-1.53	-1.98	1.25
School B1	M	6–8						
School B2	M	5–8						
Group Δ			16	-5.28	-9.14	-5.15	-5.06	7.88
School E1	M	6–8						
School E2	M	5–8						
Group Δ			4	6.77	4.63	-3.44	0.96	-3.57
School H1	M	6–8						
School H2	M	6–8						
Group Δ			**-172**	**11.07**	**10.51**	3.3	6.27	0.85
School O1	S	9–12						
School O2	S	9–12						
Group Δ			79	-0.22	3.63	-8.44	-2.12	-6.2
School L1	M	PK–8						
School L2	M	PK–8						
Group Δ			-51	3.77	-1.31	2.68	-4.24	1.72

Table 9 continued next page

Table 9. Summary of School Matching Results, continued

School Name	Type	Grade Span	Enroll-ment	FRPL %[1]	Diversity %	Sp Ed %[2]	Mathe-matics % 2010[3]	Reading/ Language Arts % 2010[3]
School J1	E	PK–5						
School J2	E	PK–6						
Group Δ			-11	6.27	-1.4	5.07	-1.22	-2.09
School C1	M	6–8						
School C2	M	5–8						
Group Δ			-127	-4.88	1.89	-4.84	-6.01	-2.84
School A1	M	6–8						
School A2	M	PK–8						
Group Δ			45	1.56	-4.52	2.88	-6.23	-2.61
School G1	S	6–12						
School G2	S	6–12						
Group Δ			218	26.22	11.32	11.84	-0.52	-7.69
School D1	M	PK–8						
School D2	M	PK–8						
Group Δ			-38	-4.47	0.61	-1.48	3.4	3.35
School F1	E	PK–5						
School F2	E	PK–5						
Group Δ			41	-5.8	-3.82	6.68	-3.94	1.12

*In order to protect the identify of these schools the scores have been redacted.

[1] = Free and Reduced Price Lunch Eligibility

[2] = Special Education Eligibility

[3] = SY2010 Mathematics and Reading/Language Arts Proficiency Rate

Our research design (described below) for analysis at the program level used panel data, which required including only those students in our model who had two consecutive test scores at consecutive grade levels. After removing students who did not meet these two criteria, each group (SIG schools and non-SIG comparison sample) included approximately 3,000 students for a total of 5,924 students (see Table 7, page 14, for a breakdown by grade level).

Confirmation of matching process

We began by testing for the presence of statistically significant differences between our two groups in baseline achievement (i.e., SY2010). We assumed that, if no differences were found between them at baseline, our groups were reasonably matched on achievement prior to the implementation of the SIG program. This was a critical assumption that we believed must be met to ensure that subsequent comparisons of achievement trends during

the implementation period were made on a fair basis. We conducted a series 16 independent *t* tests to test this assumption (one per content area and grade), using group membership as the independent variable and SY2010 scale scores as the dependent variable.

The tests for mathematics provided robust evidence that the matching process was successful, with the exception of Grade 8. Tests for reading/language arts provided evidence that the matching process was successful for Grades 3–6 and 10, but not for Grades 7–9. As a result of the failure to match adequately in Grade 8 for mathematics and Grades 7–9 in reading/language arts, we do not present the results of RM ANOVAs for these grade levels in the results section. However, we do present descriptive information with the caveat that these data should *not* be used to judge the quality of the SIG program. Table 10 and Table 11 provide an overview of the results of the *t* tests. The final column indicates whether or not we deemed the match to be successful for that grade level/content area.

Table 10. SY2010 Mathematics Matching Confirmation Pre-RM ANOVA

	Mathematics Scale Score for SY2010								
	SIG		NON-SIG					Mean	Matching
Grade	Score	SD	Score	SD	*t*	*df*	*p*	Δ	success
3	558.39	39.59	560.54	38.49	-.690	627	.491	-2.155	YES
4	568.17	48.21	563.77	49.55	.985	494	.325	4.396	YES
5	587.12	44.02	588.75	51.44	-.226	381	.821	-1.627	YES
6	608.01	43.94	609.30	44.47	-.561	1485	.575	-1.288	YES
7	622.50	51.15	625.44	51.64	-1.08	1445	.279	-2.942	YES
8	622.19	55.59	600.67	54.298	2.46	163	.015	21.525	NO
9	631.20	59.62	627.31	55.81	.882	688	.378	3.887	YES
10	632.00	50.02	633.14	51.03	-.276	596	.783	-1.142	YES

Table 11. SY2010 Reading/Language Arts Matching Confirmation pre-RM ANOVA

	Reading/Language Arts Scale Score for SY2010								
	SIG		NON-SIG					Mean	Matching
Grade	Score	SD	Score	SD	*t*	*df*	*p*	Δ	success
3	418.40	41.59	419.09	40.73	-.209	627	.834	-.689	YES
4	426.93	50.06	425.14	40.28	.439	494	.661	1.782	YES
5	440.21	38.70	440.46	42.89	-.043	381	.966	-.258	YES
6	450.21	45.33	454.09	43.17	-1.684	1485	.092	-3.881	YES
7	459.64	41.05	464.40	37.93	-2.263	1445	.024	-4.755	NO
8	448.82	38.85	468.19	39.42	-3.071	158	.003	-19.36	NO
9	455.00	59.22	466.37	45.32	-2.817	688	.005	-11.377	NO
10	450.29	63.36	458.55	60.48	-1.616	593	.107	-8.257	YES

Measures and covariates

Test scores analyzed were generated using West Virginia's summative assessment, WESTEST 2. Only mathematics and reading/language arts scores were used in this analysis.

Research design

Program level investigation

To analyze program effects, we collected available testing records for all students who were enrolled in the SIG and non-SIG comparison schools during SY2009, SY2010, and SY2011. For the main analysis, we used SY2010 data as the baseline point, given that it was the most recent testing point prior to SIG implementation. The longer term plan is to examine achievement trends over the entire course of the SIG program implementation period, which will ultimately include SY2011 and SY2012; however, for this report, we had access only to SY2011 data as a postimplementation data point[6].

Once *t* tests indicated matching was successful, we conducted RM ANOVAs using panel data. Our models included individual students who progressed up the grade span across 2 years of time (e.g., Grades 3–4, Grades 4–5, etc.) rather than examining different cohorts of students over time (e.g., comparing those students enrolled in Grade 3 during the 2010 school year [SY2010] to those enrolled in Grade 3 during [SY2011])[7]. It should also be noted that our ultimate analyses compared achievement gains between the *group* of students in SIG schools and the *group* of students in matched non-SIG comparison schools. RM ANOVAs were not conducted at the school level. Rather, we compared all students in SIG schools to all students in matched non-SIG comparison schools within each grade level cohort to examine the impact of the SIG program statewide, not to attempt to evaluate individual implementations of the SIG program, which differ by school.

We used the RM ANOVAs to detect the presence of any statistically significant interaction effects between time and group membership (SIG or non-SIG comparison) on mathematics and reading/language arts achievement. The presence of statistically significant interactions indicated that one group of students outperformed the other over time by a significant margin. In these cases, follow-up descriptive analyses examined which group of students (i.e., SIG or non-SIG comparison) exhibited higher performance over time.

School level investigation

Because we conducted the RM ANOVAs to examine program-level differences (i.e., SIG versus non-SIG comparison), the results did not provide substantive information about the efficacy of the 14 individual school-level SIG implementations in Cohort 1. To begin to address this shortcoming, we developed a descriptive school-level achievement profile for each Cohort 1 SIG school. Our goal was to provide descriptive information that Title I staff and each school could use to begin developing an understanding of how the reading/

[6] The final report will include data for SY2010, SY2011, and SY2012.

[7] We did not include students who were retained or who did not take WESTEST 2 in both years in our models.

language arts and mathematics proficiency gains that have occurred at the conclusion of the first SIG implementation year relate to those gains exhibited in reasonably comparable schools that are not receiving SIG resources[8]. The profiles are preliminary given that the SIG program implementation period is not yet complete.

Findings from the program-level and school-level analyses can be found in the Results section, beginning on page 44.

Evaluation Question 4

EQ4. To what extent have SIG schools successfully implemented and institutionalized practices and structures that are supportive of continuous school improvement?

To investigate this question, we analyzed survey data collected using the Measure of School Capacity for Improvement (MSCI), and data collected onsite by Office of Title I program staff using the School Culture Survey (Valentine, 1998; Quinn, 2008) during the 2009-2010 and 2010-2011 school years.

Participant characteristics

All available professional staff from 14 of the original 15 Cohort 1 SIG schools participated in the MSCI and School Culture Surveys. One school, Mount Hope High School, closed in June 2011, so it is omitted from this report.

Sampling procedures

Professional staff from the full set of Cohort 1 schools was included in these two components of the study; thus no sampling procedures were needed.

Measures and covariates

Measure of School Capacity for Improvement

The MSCI is an instrument with demonstrated reliability and validity that schools can use to help identify their strengths and weaknesses, in order to accurately target school improvement and professional development goals (Hughes, et al., 2006). The instrument is made up of 58 items with a 6-point Likert-type scale and measures the perceptions of professional staff members on how they believe their school is faring in seven areas related to capacity for improvement. The seven dimensions of capacity for improvement are described below.

1. *Equity in practice* assesses "…equitable practices in the school—specifically, responsive pedagogy and anti-discriminatory practices."
2. *Expectations for student performance* assesses "…staff members' expectations of the students and their beliefs that all students can perform well academically."

[8] The non-SIG comparison schools used for these analyses were those indicated earlier in this report in Table 9, page 15.

3. *Differentiated instruction* addresses "…instructional practices and strategies for reaching students of diverse learning needs."
4. *Improvement program coherence* pertains to the "extent to which improvement initiatives and efforts at a school are coordinated."
5. *Peer reviewed practice* explores "…the observation and review by staff of their peers' work."
6. *Coordinated curriculum* addresses "…the coordination of curriculum within and across grade levels at the school."
7. *Technical resources* concerns "instructional resources and materials, including whether staff possess or have immediate access to adequate materials and resources to achieve instructional objectives." (Hughes, et al., 2006, pp. 2–3).

School Culture Survey

The School Culture Survey measures perceptions of professional staff regarding school culture. The instrument is made up of seven domains with a total of 45 items. Six of the domains (35 items) were identified by Gruenert and Valentine (1998) while the seventh domain (10 items) was adapted from David Quinn (2008). The seven domains are described below.

1. *Collaborative leadership* describes the degree to which school leaders establish and maintain collaborative relationships with school staff (11 items).
2. *Teacher collaboration* describes the degree to which teachers engage in constructive dialogue that furthers the educational vision of the school (6 items).
3. *Professional development* describes the degree to which teachers value continuous personal development and school-wide improvement (5 items).
4. *Unity of purpose* describes the degree to which teachers work toward a common mission for the school (5 items).
5. *Collegial support* describes the degree to which teachers work together effectively (4 items).
6. *Learning partnership* describes the degree to which teachers, parents and students work together for the common good of the student (4 items).
7. *Efficacy factor* assesses the extent to which teachers consider their students to be capable of successful learning and consider themselves to be responsible and effective agents in instructing students (10 items).

Research design

Measure of School Capacity for Improvement

The pencil/paper survey was distributed to school improvement specialists (SISs) in mid-May to administer the MSCI for completion by all professional staff at all SIG schools (N=759). For schools without a school improvement specialist, the LEA Title I director was

charged with the responsibility of administering the survey. The survey was completed by faculty and staff members (n = 571; 75.2% response rate) before the end of school year 2010-2011 in all 14 SIG schools of various programmatic levels.

School Culture Survey data collection and analysis

The majority of data for the survey was collected by program staff during school year 2010-2011, with the exception of Guyan Valley Middle and Hamlin PK-8 (Table 1). For these two schools, data were not available for the 2010-2011 school year, therefore, this report uses data collected during the 2009-2010 school year. It should be noted that all data used for analysis in this report come from the first school culture survey administered in SIG schools and, as such, these data are considered baseline. As much as possible, researchers attempted to narrow down the time of data collection, because the survey should be administered around the same time annually. However, with the exception of four schools, we were not able to ascertain the month during which the survey was administered. Dates with asterisks in Table 12 represent the time of data entry into the online system or what was verbally communicated to us by program staff. It should also be noted that since the potential pool of survey participants is dependent on the size of each SIG school, the number of respondents also vary by school, ranging between 15 and 43 for an average of 28 respondents per school (Table 12). In all there were 396 respondents.

Table 12. Time of baseline data collection and number of respondents

School	Time of data collection	Number of respondents
Cedar Grove Middle	Feb-2010	19
Doddridge County Elementary	Jan-2010	15
East Bank Middle	Dec-2010 *	28
Franklin Elementary Center	Jan-2010	27
Geary Elementary/Middle	2010 **	29
Guyan Valley Middle	2009 **	24
Hamlin PK-8	2010 **	25
Malden Elementary	Sept-2010 *	15
Martinsburg North Middle	Sept-2010	37
Mount View High	Oct-2010 *	43
Riverside High	Nov-2010 *	42
Sandy River Middle	Oct-2010 *	21
Stonewall Jackson Middle	Dec-2010 *	42
West Hamlin Elementary	2009 **	29

*Time of data entry
**Based on verbal communication

We used Pearson's correlation coefficient (r) to measure the strength of relationships between the School Culture Survey data and proficiency rates on the four content areas tested by the WESTEST 2 for SIG schools. First, we tested the strength of relationship between school culture survey mean scores and proficiency rates on WESTEST 2 for the four content areas. We then tested the strength of relationship between each of the seven domains that make up the school culture survey and proficiency rates on the four content areas. Descriptive statistics were also interpreted for each of the seven domains at the program level (aggregated across all SIG schools) and at the individual school level. Findings from these analyses can be found in the Results section, beginning on page 81.

Results

Results of the various investigations included in this evaluation study are organized by evaluation question (EQ), beginning with EQ1.

Evaluation Question 1

EQ1. To what extent has the SEA provided adequate support to LEAs and participating schools to successfully lead school improvement efforts?

To investigate this question, we reviewed documents associated with the School Improvement Grant (SIG) program and conducted an End-of-Year Survey of educators working in SIG schools. For a full description of the methods used, see page 7.

Document review

The document review yielded information about the number of professional develop sessions offered and the level of participation in them, and the number and nature of technical assistance visits reported.

Professional development sessions

Four major WVDE-sponsored professional development sessions (School Leadership Team Conferences) were provided during school year 2010–2011 (Table 13), which were attended by a total of 425 participants, including representatives from 11 Tier I state schools. The conference in October 2010 had the largest attendance with 148 participants, while the one held in August 2010 for school improvement specialists and Title I directors had the smallest attendance with 39 participants. Professional development sessions in February 2011 and April 2011 had 120 and 118 participants respectively. For themes of professional development sessions and characteristics of participants please refer to Table 13 below.

Table 13. Characteristics of Professional Development Sessions and Participants

Date of PD sessions	Location of PD sessions	Theme	Participant role group	Number of participants*
Total				425
August 3-4, 2010	Bridgeport Conference Center	Roles and responsibilities of School Improvement Specialists	School Improvement Specialists and Title I Directors	39
October 18-29, 2010	Bridgeport Conference Center	School Culture and Data Utilization	LEA contacts and School Leadership Team members	148
February 16, 2011	Bridgeport Conference Center	Collaborative Teams and Curriculum (CSOs)	LEA contacts and School Leadership Team members	120
April 7, 2011	Embassy Suites in Charleston	Collaborative Teams and Student Leadership	LEA contacts and School Leadership Team members	118

*Includes representatives from 11 Tier I state schools

Technical Assistance Events

There were a total of 358 technical assistance reports posted on the WVDE Statewide System of Support (SSOS) websites for the 14 SIG schools by the WVDE Title I coordinators and school improvement specialists who, among others, were chiefly responsible for providing support to the schools. These reports documented the name and title of the technical assistance provider, county and school names, date of the visit, and summary of assistance provided. The reports also indicated to whom reports had been sent. In most cases reports were sent to school principals, district personnel assigned to school, state school improvement specialist or Title I liaison, and relevant WVDE personnel. While the nature of the technical assistance provider visits varied, the vast majority involved assistance related to fostering collaboration, strategic planning, and the collection and utilization of various types of data (e.g., Acuity, Instructional Practices Inventory, School Culture Survey) to drive school improvement efforts. Table 14 displays the number of technical assistance reports posted on the SSOS websites for each SIG school.

Table 14. Number of Technical Assistance (TA) Reports by SIG School

School	Number of TA reports*
Total	358
Martinsburg North Middle	46
Doddridge County Elementary	13
Cedar Grove Middle	43
East Bank Middle	48
Malden Elementary	29
Riverside High	42
Stonewall Jackson Middle	39
Guyan Valley Middle	41
Hamlin PK-8	7
West Hamlin Elementary	3
Mount View High	6
Sandy River Middle	7
Geary Elementary/Middle	27
Franklin Elementary Center	7

*Some reports indicate TA was provided to more than one school per visit. The count, therefore, does not account for the duplication of records.

End-of-Year Survey results

The online End-of-Year Survey was completed by seven Title I directors and one transformation specialist. As described in the Methods section (page 8), the survey was intended to elicit feedback regarding the quality, relevance, and usefulness of professional development and technical assistance provided by WVDE to schools and districts, and to schools by their associated district personnel.

Perceived quality and impact of professional development

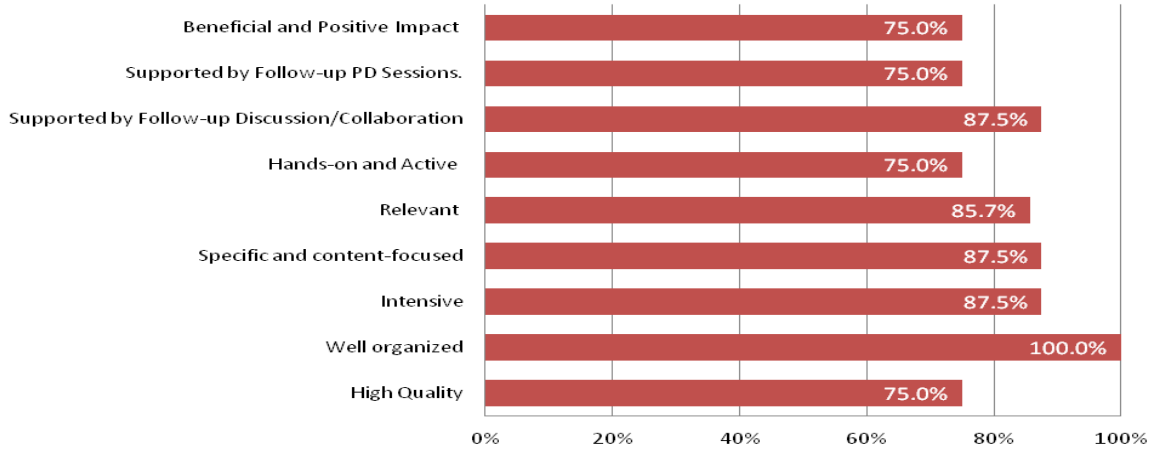
Participants' responses about the quality of the professional development and its main components are provided in Table A- 1 (page 117 in Appendix A), Table A- 2 (page 118), and Table A- 3 (page 118) as well as Figure 1 (page 25), Figure 2 (page 26), and Figure 3 (page 26). Respondents were asked to respond to each statement using a 5-point Likert-type response format (1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Neutral*, 4 = *Agree*, 5 = *Strongly agree*). Mean scores were calculated for each item by summing respondents' ratings and dividing the resulting value by the total number of responses for that item. The count and percentage of participants' responses for each survey item are summarized in the three tables along with the average ratings. The figures show the percentage of favorable responses for each survey item. Percentages of favorable responses were calculated for each item by summing the number of responses of *Agree* and *Strongly Agree* and dividing that value by the total number of responses for that item.

The midpoint of the scale, a *Neutral* response (3.00), is used as a point of reference to interpret data from the survey. Thus, any mean score below 3.00 would indicate general disagreement and any mean score above 3.00 would indicate general agreement with an item. A higher mean score (i.e., 4.00 or above) indicates that the professional development events were perceived to be of higher quality.

Quality of professional development

Participants' responses to nine items about the quality of the professional development events they attended are summarized in Table A- 1 (page 117) and Figure 1. The overall mean for the nine items was above the threshold of agreement (M = 4.30; SD = 0.14). All respondents (n = 8) either agreed or strongly agreed that professional development sessions were well organized, making this item the highest rated (M = 4.50; SD = 0.27). The extent to which professional development sessions were hands-on and included active learning opportunities received the lowest average rating (M = 4.00; SD = 0.17). It should be noted here that this item was still above the threshold of agreement. Participants also rated the professional development sessions they attended as high quality (M = 4.25; SD = 0.21); intensive in nature (M = 4.25; SD = 0.23); relevant (M = 4.29; SD = 0.24); supported by follow-up sessions or collaboration at school, office, or online (M = 4.38; SD = 0.23); supported by related follow-up professional development sessions (M = 4.25; SD = 0.21); and beneficial, having a positive impact on students and/or school (M = 4.38; SD = 0.26). At least six out of eight respondents (75%) had favorable responses to each of the nine items in this section of the survey.

Figure 1. End-of-Year Survey Respondents' Perception of Quality of Professional Development

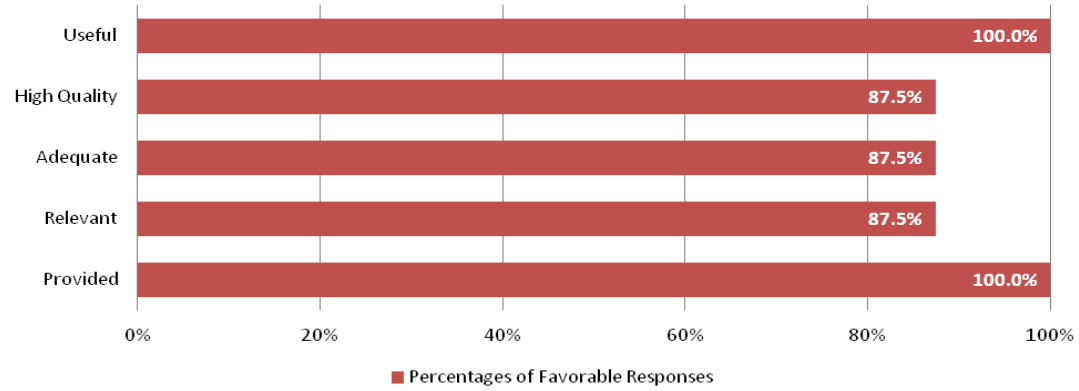


Quality of training materials and resources

Survey respondents were asked to respond to five items about training materials/resources provided during professional development sessions. The average ratings for each item are presented in Table A- 2 (page 118) along with the overall mean score for this component of the professional development sessions. The overall mean rating for training materials/resources was above the threshold of agreement (M = 4.33; SD = 0.17). All respondents (100%) agreed or strongly agreed that materials/resources were provided during

training sessions, making this survey item the highest rated, with a mean score of 4.63 (SD = 0.29). Respondents also indicated that the materials/resources were relevant (M = 4.25; SD = 0.23), adequate (M = 4.25; SD = 0.23), high quality (M = 4.25; SD = 0.23) and useful to their work (M = 4.25; SD = 0.23). At least 87.5% of respondents (n = 7) had favorable responses to each of the five items about the quality of training materials/resources (Figure 2).

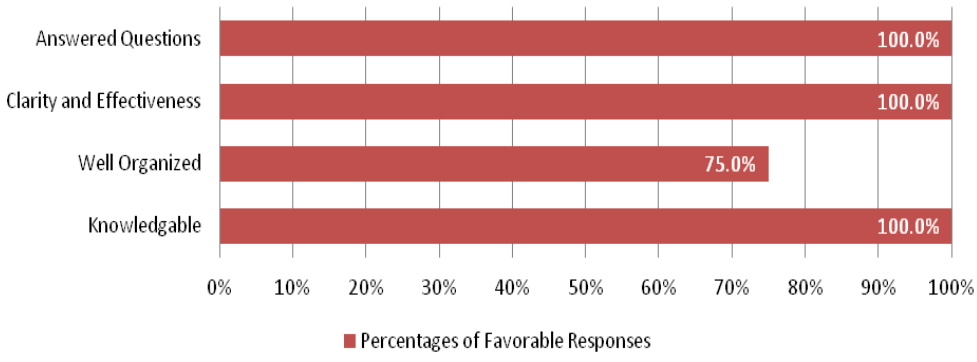
Figure 2. End-or-Year Survey Respondents' Perception of Quality of Training Materials/Resources



Quality of presenters

Respondents were provided with four statements about presenters at professional development sessions and were asked to indicate their level of agreement or disagreement with the statements using the 5-point Likert-type response format. The overall mean rating for the presenters was 4.32 (SD = 0.12), which was above the level of agreement (Table A- 3, page 118). The average rating for the degree to which participants perceived presenters to be well organized was the lowest among the four items but still above the threshold of agreement (M = 4.13; SD = 0.12). Six out of eight (75%) of respondents had favorable responses regarding this survey item (Figure 3). The remaining three items (knowledge of the topic; clarity and effectiveness of presentation; and the degree to which presenters answered questions adequately during sessions) were rated equally with a mean rating of 4.38 (SD = 0.29) and 100% of respondents (n = 8) either agreed or strongly agreed with these three survey items.

Figure 3. End-or-Year Survey Respondents' Perception of Quality of Presenters



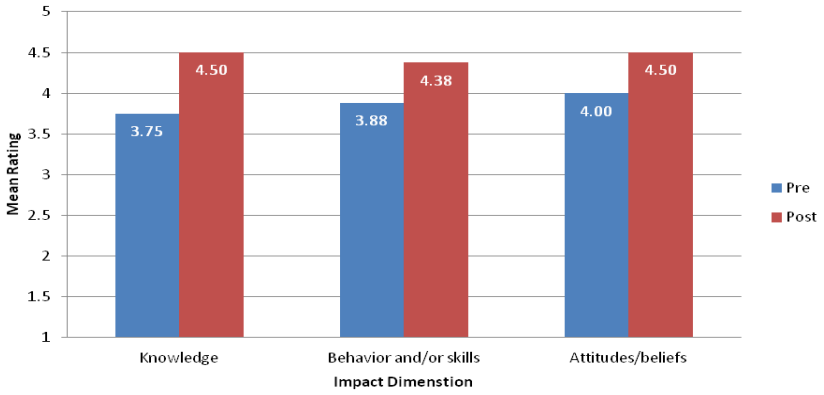
Impact of professional development sessions

The End-of-Year Survey included a retrospective pretest/posttest component to assess the extent to which respondents perceived a change in their own knowledge, behaviors, and beliefs/attitudes. Respondents were provided with three pairs of items and were asked to rate the degree to which they agreed with statements about themselves both before and after having participated in the professional development sessions using a 5-point Likert-type response format.

It should be noted that retrospective pretest/posttest items have some limitations or threats to validity. For example, participants may not accurately recall knowledge, behaviors, and beliefs/attitudes held in the past. On the other hand, respondents may indicate change or improvement to fit their own expectation or that of the professional development session even if it did not occur.

In all cases, survey respondents rated themselves higher after participating in professional development sessions than before participating (Table A- 4, page 119; Figure 4). Pretest/posttest results indicate that the greatest amount of gain was in acquiring knowledge about the various topics covered in the four major professional developments with a mean difference of 0.75. The gains for behavioral and attitudinal change were rated equally with a mean difference of 0.50. NOTE: The sample size (n = 8) was too small to run statistical significance tests.

Figure 4. End-of-Year Survey Respondents' Perception of the Impact of Professional Development (Pre-/Postsession)



Participant comments about the quality and impact of professional development

Each section of the survey included at least one open-ended question that provided respondents an opportunity to give feedback in their own words (qualitative data). Four participants (50%) provided comments about the school improvement professional development sessions they attended. The comments substantiate the positive findings of the quantitative data.

One respondent indicated the sessions served as confirmation that his/her knowledge, skills, and work aligned with current best practice in the area of school improvement. For two other respondents the sessions were helpful in making "connections"

with what they have learned in the past, and more specifically, with contents covered at earlier professional development sessions. However, one of the two participants indicated that those "connections" were not made as evident during the professional development session held in April as compared with the other three.

One respondent indicated that the professional development sessions were practical and paced in a way that allowed the school and district leadership teams to discuss and implement what they learned. Another respondent found the supporting materials provided during the sessions to be helpful.

Perceived quality of the statewide system of support teleconferences

The purposes of the statewide system of support (SSOS) teleconferences, also known as Collaborative Conversations, are to provide leadership teams from each Tier I and SIG school (a total of 27 schools) the opportunity to discuss current and future improvement efforts, and enable SSOS personnel (e.g., WVDE and RESA representatives) to better focus support provided to schools. Local participants in these conversations included principals, assistant principals, lead school staff members, the superintendent, and school improvement/transformational specialists.

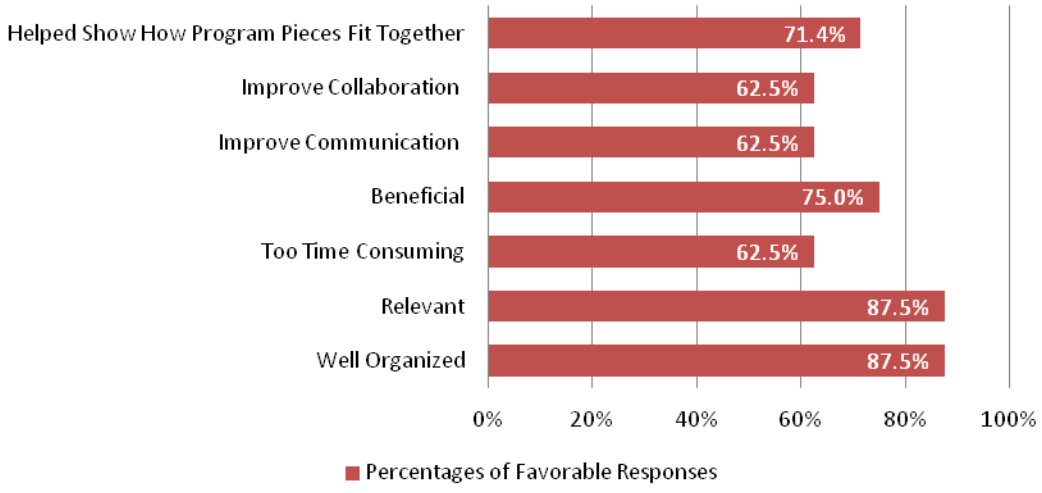
Each Tier I and SIG school participated in these conversations three times during the 2010-2011 school year. Prior to each of the three teleconferences schools received a set of general questions intended to guide the conversations. During the teleconferences, school leadership teams were given 15 minutes to address the questions they received prior to the teleconference. Following the school response, members of the support team for each school (e.g., school improvement specialists and RESA staff) were given 10 minutes to present their perspective on the school's improvement process. An additional 5 minutes was then provided for a perspective from the superintendent. A few days after the teleconference each school received a letter containing suggested "points for discussions" from the Office of School Improvement, intended to assist school and district leadership teams in their on-going school improvement efforts.

In seven survey items, survey participants were asked about the quality of the SSOS teleconferences. NOTE: Unlike previous survey items which used a 5-point Likert-type response format to calculate the mean, survey items in the sections below required participants to respond to each statement using a 4-point Likert-type response format (1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Agree*, and 4 = *Strongly agree*). A fifth response category, *Not applicable*, was included to give respondents the option to pass on a question, if any of the items did not apply to them. The *Not applicable* responses were not used to calculate the mean scores. Thus, a higher mean score (i.e., 3.00 or above) for each survey item indicates a favorable response to each statement. Response categories for the only negatively worded survey item in this section—the degree to which teleconferences were too time consuming—were assigned the following numerical values: 1 = *Strongly Agree*, 2 = *Agree*, 3 = *Disagree*, and 4 = *Strongly disagree*.

The overall mean rating for the quality of the teleconferences was 2.94 (SD = 0.12) (Table A- 5, page 120). Three out of eight respondents (37.5%) either agreed or strongly agreed that the teleconferences were too time consuming, making this item the lowest rated

among the seven survey items regarding the teleconferences (M = 2.38; SD = 0.27). Two additional items also had mean scores below 3.00: the degree to which the teleconferences helped respondents see how the various pieces of the SIG program fit together received (M = 2.86; SD = 0.24) and the degree to which the teleconferences helped to improve collaboration between the different levels of the school system (M = 2.88; SD = 0.18). It should be noted here that even though the mean rating for the three items discussed above is below 3.00, at least 62.5% of responses to each item were positive or favorable (Figure 5). On the remaining four items, respondents indicated that the conferences were well organized (M = 3.25; SD = 0.23), relevant to their work (M = 3.25; SD = 0.23), have served as avenues of beneficial feedback (M = 3.00; SD = 0.20), and have helped to improve communication between the different levels of the school system (M = 3.00; SD = 0.18). At least 70% of those who responded to these four items had favorable responses regarding the SSOS teleconferences.

Figure 5. End-of Year Survey Respondents' Perception of the Quality of SSOS Teleconferences



Participant comments about the quality of the SSOS teleconference

Participants were asked three specific open-ended questions: (a) what they liked *best* about the teleconferences; (b) what they liked *least* about the teleconferences; and (c) if they could change something about the teleconferences, what it would be. A fourth question provided respondents an opportunity to offer any additional comments.

Six respondents (75%) provided comments regarding what they liked *best* about the teleconferences. Comments from two participants indicated that they liked the constant feedback and communication among stakeholders, including input from teachers and superintendents. Feedback from another respondent indicated that the teleconferences "forced schools to stay focused on some specific school improvement areas," which this respondent considered to be a positive attribute. That schools answer a predetermined set of questions and the preparation they put into each one of the collaborative calls were also mentioned as positive aspects. One respondent stated that there was nothing likable about the teleconferences.

Six of the eight respondents (75%) provided comments regarding what they liked *least* about the teleconferences. One of the six respondents didn't have a negative comment about the collaborative calls. Two respondents indicated that conversations tended to "skew" toward topics or areas that were outside of their current work. Both respondents claimed the topics of conversations were more aligned with the interest areas of participants from the state department and in the case of one of the respondents this was a cause of distraction. Two other respondents commented that participants in the teleconferences did not appear to be honest or forthcoming, with one of the respondents describing the conversation as "contrite, staged, and rehearsed." It was not clear from the comment of one of these respondents who the remark was directed to. The other respondent, however, stated that initially schools were "trying to make things seem better than what they are" but toward the end of the conversations schools seemed "to realize [they needed] to be more open and honest." One respondent liked least the fact that schools did not discuss real time data and what they were doing about it.

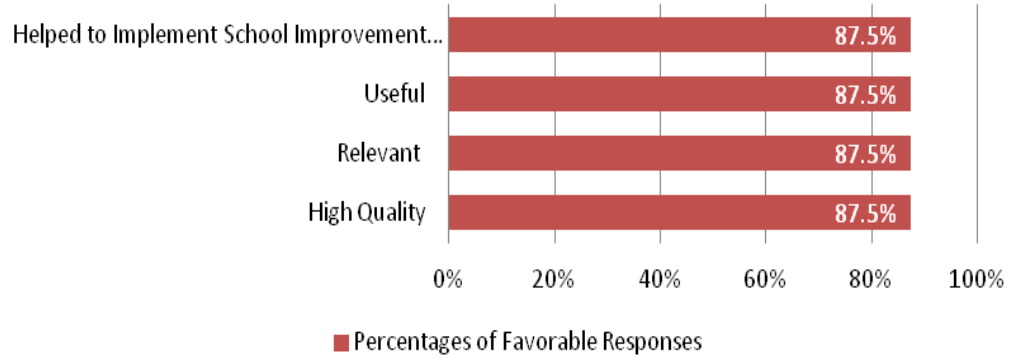
Five participants (62.5%) responded to the item about what they would *change* about the teleconferences. One respondent did not wish to change anything, while another respondent said to "eliminate" the teleconferences altogether but did not elaborate any further. One comment indicated a preference for the same state department representatives to participate in all the conversations throughout the year. A respondent commented that the process should allow all school team members to answer questions, while another thought the process was rushed and wanted more time to "delve into meaningful conversations that would include feedback, reflection, and setting more specific action steps as a whole team."

Only two participants (25%) responded to the last open-ended question (any additional comments) regarding the SSOS teleconferences. One expressed the desire to be notified beforehand if questions will be used as guidelines for the May presentations. The other respondent pointed out that the monthly meetings with the transformation specialists have been helpful to their system.

Perceived quality of technical assistance

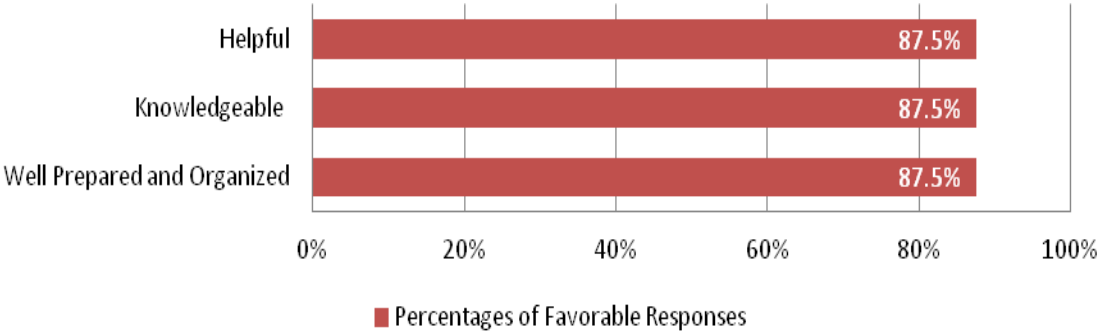
Respondents were provided with four statements about technical assistance that SIG schools in their district received from the WVDE Office of Title I during the school year 2010-2011. They were asked to indicate the level with which they agreed or disagreed with the statements. The overall mean rating for the quality of technical assistance was 3.35 (Table A- 6, page 121). There was very little variation in mean rating among the four survey items (SD = 0.06) and 87.5% of the eight respondents (n = 7) either agreed or strongly agreed with all of the statements regarding the quality of technical assistance (Figure 6). Respondents indicated that the technical assistance was high quality (M = 3.25; SD = 0.23), relevant (M = 3.38; SD = 0.27), useful (M = 3.38; SD = 0.27), and helped to implement school improvement efforts (M = 3.38; SD = 0.27).

Figure 6. End-of Year Survey Respondents' Perception of the Quality of Technical Assistance



Respondents were asked to rate the technical assistance provider and were presented with three survey items to rate using the 4-point Likert-type response format. The overall mean rating for the quality of technical assistance was 3.42 and there was very little variation in mean rating among the three survey items (SD = 0.07) (Table A- 7, page 121). The majority of respondents, seven out of eight (87.5%), either agreed or strongly agreed with all survey items (Figure 7). They indicated that the provider was well prepared (M = 3.38; 0.27), was knowledgeable about the topic (M = 3.38; 0.27), and was helpful (M = 3.38; 0.27).

Figure 7. End-of Year Survey Respondents' Perception of the Quality of the Technical Assistance Provider



Participant comments about the quality of the technical assistance and its provider

Six participants (75%) responded to an item that asked participants to provide any additional comments about technical assistance they have received from WVDE Office of Title I. The qualitative feedback received from the six respondents substantiates the strongly positive results from quantitative data presented above.

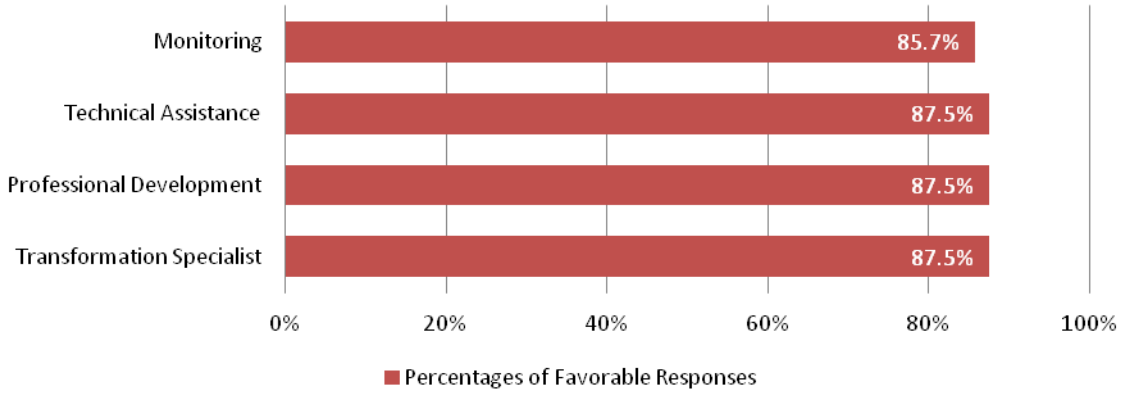
Overall, technical assistance providers were perceived to be very knowledgeable, committed, accessible, and extremely helpful. Two respondents commented on their providers' ability to build positive relationships with the staff, which has facilitated the implementation of the SIG and promotion of change. Respondents indicated that technical assistance

providers were instrumental in achieving the success experienced by SIG schools thus far. One respondent specifically stated that individual visits by the technical assistance providers were "much more beneficial than the PD sessions."

Perception of benefit to schools from services provided by districts

The survey included items designed to gather data on the types of services districts provided to SIG schools. Respondents were asked if the district had provided any of the following to SIG schools during the 2010-2011 school year: (a) a transformation specialist; (b) professional development supported by SIG funds; (c) technical assistance related specifically to implementation of the SIG; and (d) monitored implementation of the SIG. All respondents answered *Yes* to this question (n = 8). Participants were then provided with four statements regarding the various services provided by districts to SIG schools and were asked to indicate their perception of the degree to which SIG schools have benefited from each of the services. The overall mean score for all items was 3.29 (Table A- 8, page 122). There was very little variation between mean scores among the four items, ranging from 3.25 to 3.38 (SD = 0.06). At least 85.7% of respondents either agreed or strongly agreed that school benefited from all four types of services provided by districts (Figure 8).

Figure 8. End-of Year Survey Respondents' Perception of the Benefit to SIG Schools from Services Provided by Districts



Participant comments about the benefit to SIG schools from services provided by districts

Five participants (62.5%) responded to an item which asked participants to provide any additional comments they may have about services the district has provided to SIG schools during the 2010-2011 school year. One respondent emphasized the importance of support and oversight provided to schools while another respondent indicated that services provided by the district were focused to the needs of each school. Comments from two of the respondents highlight the importance of a full-time transformation specialist devoted to individual schools and another respondent stated that weekly visits from the district director seem to have a positive effect on the school's efforts.

Additional comments

In addition to the opportunity to provide qualitative feedback at the end of each survey section, respondents were asked three general open-ended questions to conclude the survey. They were asked to share (a) challenges or barriers they have encountered in utilizing/implementing what they have learned, (b) additional resources/support that would be useful, and (c) any additional comments.

Challenges/barriers. Six participants (75%) responded to the item concerning challenges or barriers to the utilization/implementation of what districts have learned at professional development sessions. Two of the respondents indicated that challenges centered on *defining roles*. One of these respondents indicated the challenge was defining the role of the school leadership. The other respondent, on the other hand, commented that barriers were rooted in defining the roles of SIG staff which created "difficulties for the teaching staff not being clear on how to function with the SIG staff." The respondent further elaborated by stating that they [county] were, in time, able to clarify roles and responsibilities which led them to move forward in the "component of the SIG."

Two other respondents identified *time* as a challenge or barrier that districts have encountered. One respondent indicated that school staff spend too much time away from school in activities related to SIG and that these activities should take place at the building level. The other respondent stated that although they have utilized the ideas and strategies they have learned at state meetings, a change of administration mid school year has presented time challenges. The respondent did indicate, however, that the new administrator is "strong" and the school is "making progress".

The other two respondents identified *communication* and *unwillingness* of some staff to change as on-going challenges.

Resources/support. Five participants (62.5%) provided feedback related to additional resources/support that would be useful to successfully implement what they learned during the professional development sessions. One respondent indicated that adequate amounts or resources/support have been provided. Assistance with budgets and expenditures was identified by one respondent while another respondent pointed to staff members who "refuse to change" as a challenge to bring about "positive culture change." The latter comment indicates an area where additional support may be targeted.

One respondent indicated a preference to have a portion of each school improvement networking session devoted to a particular element of the SIG. The respondent believes that this would allow districts to "see where we [district] should realistically be, get ideas of best practice in those areas, and share concerns and successes in those areas." Another respondent stated that since the data show that the strategies they are utilizing are working, what would be most helpful is to "continue with the same themes/initiatives for the rest of the grant to give them more time to work as long as the data shows continuous improvement."

Additional comments. Only three participants (37.5%) provided feedback to the last item asking for any additional comments. One respondent indicated that they have greatly benefited from the work in the SIG school. The other two respondents reiterated what they have previously stated. One was regarding concern about the amount of time staff is away

from the building, asking "how can schools improve when the best teachers and principal are away?" The other respondent emphasized the importance of clearly defined roles and responsibilities of stakeholders in the school improvement effort.

Evaluation Question 2

EQ2. To what extent has the SIG program resulted in improved quality and quantity of instruction and nonacademic support?

To investigate this question, we used two approaches: (a) we analyzed data that were collected by Title I staff using the Instructional Practices Inventory (IPI) and (b) we conducted the Effective School Practice Survey. Results from both datasets are described below.

Instructional Practices Inventory results

As mentioned earlier (page 9), the IPI uses the following three broad categories and six coding categories:

Student engaged instruction
- (Category 6) Student active engaged learning
- (Category 5) Student learning conversations

Teacher-directed instruction
- (Category 4) Teacher-led instruction
- (Category 3) Student work with teacher engaged

Disengagement
- (Category 2) Student work with teacher not engaged
- (Category 1) Complete disengagement

We conducted two analyses using the IPI data. First, we examined the strength of the relationship between WESTEST 2 proficiency rates and IPI data for SIG schools. Analysis included various groupings of IPI categories (e.g., Categories 5 and 6, Categories 4, 5, and 6, and so on), to test where the strongest relationship exists between IPI data and proficiency rates. Second, we compared baseline IPI data from SIG schools to *typical* baseline data (Valentine, 2009).

Correlation between IPI and achievement data

We used Pearson's correlation coefficient (r) to measure the strength of the relationship between the IPI data and proficiency rates in the four content areas tested by the WESTEST 2 for SIG schools. Correlation coefficients tell us the *strength* and *direction* of a relationship between two variables. Pearson's r can vary in strength from -1 to 1. An r value of -1 or 1 indicates a perfect relationship, and a value of 0 indicates the absence of any relationship. A positive relationship indicates that if one variable increases the other variable also increases. A negative relationship indicates that when one variable increases the other variable decreases. The guideline for effect size most often used in the social sciences is that an r value between ±.1 and ±.29 indicates a weak relationship, an r value between ±.3 and ±.49 is considered a moderate relationship and an r value of ±.5 or higher indicates a strong relationship. It should be noted that regardless of the strength of the relationship one *cannot* infer a causal relationship between the two variables based on correlation coefficient values.

First, we tested the strength of the relationship between the sum of percentages of IPI Categories 5 and 6—what Collins and Valentine (Valentine, 2011) call higher-order/deeper learning—and proficiency rates on WESTEST 2 for core content areas (reading/language arts, mathematics, science, and social studies). This analysis began by adding the percentages of observations (i.e., *sum of percentages*) recorded across the SIG schools that had been coded in Categories 5 and 6. This sum of percentages was then used in the Pearson's *r* analysis for each core content area. The results showed that IPI Categories 5 and 6 have a moderate positive relationship with mathematics proficiency rates (r = .35), a weak positive relationship with proficiency rates for science (r = .22), an even weaker positive relationship with social studies proficiency rates (r = .16), and no relationship with proficiency rates for Reading/Language Arts (r = .02) (Figure 9). We also tested the strength of relationship between each IPI category and proficiency rates. Results indicate that Category 6 has a moderate positive relationship with mathematics proficiency level (r = .31) and a weak but positive relationship with science and social studies proficiency rates (r = .29, and .23, respectively). On the other hand, Category 5 showed a weak positive relationship with mathematics proficiency rates (r = .28), albeit on the high end of the weak level of relationship (Table 15). None of the correlations for either category separately or combined were statistically significant at the 0.05 level.

Figure 9. Scatter Plots of 2010-2011 Proficiency Rates and IPI Categories 5 and 6

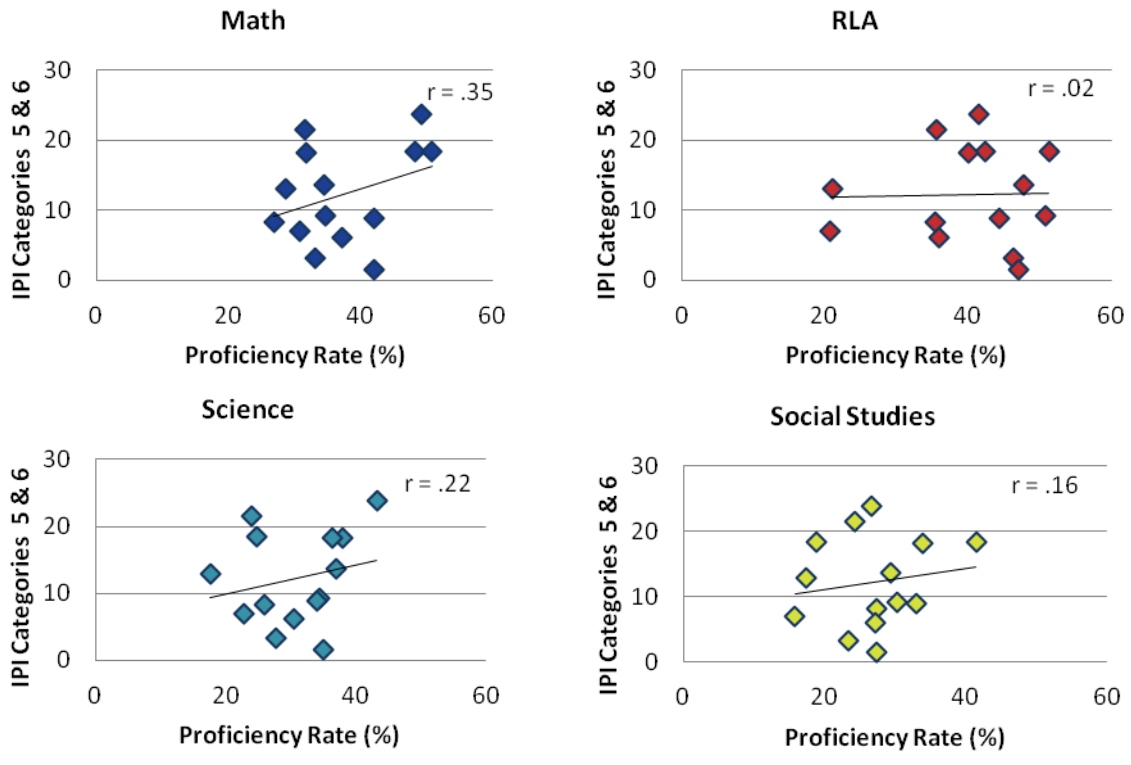


Table 15. Bivariate Correlation Between Core IPI Data and Proficiency Rates

	Correlation by core content area proficiency rate			
IPI categories	Mathematics	Reading/ language arts	Science	Social studies
6	.310	.020	.287	.233
5	.281	.025	-.045	-.085
4	.214	**.576***	.297	.194
3	-.145	**-.742****	**-.518***	-.420
2	-**.606***	.047	-.103	-.028
1	-.188	.305	.181	.209
4-6	**.519***	**.589***	**.488***	.336
1-3	-**.515***	**-.577***	**-.477***	-.327
5-6	.345	.024	.219	.163
3-4	.067	-.232	-.279	-.279
1-2	-**.458***	.260	.087	.147
2-3	-.383	**-.731****	**-.563***	-.435

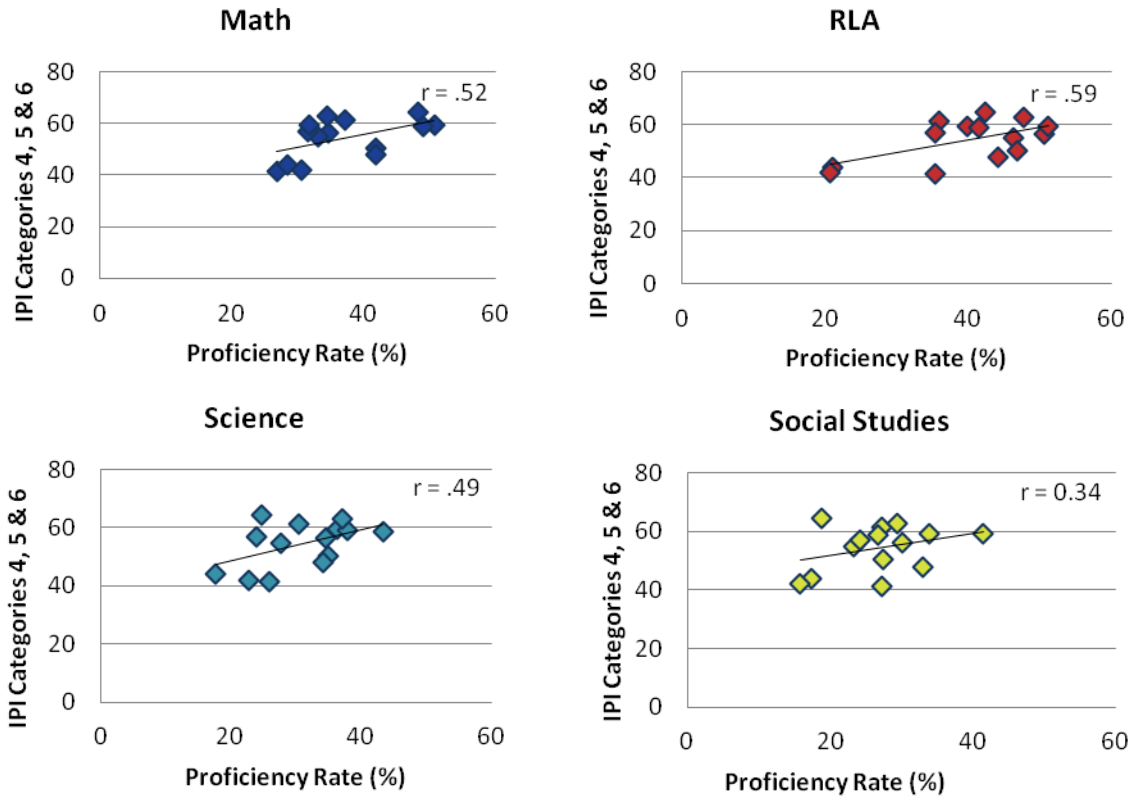
*Correlation is significant at the 0.01 level (1-tailed)

**Correlation is significant at the 0.05 level (1-tailed)

Next, we combined data for IPI Categories 4 (teacher-led instruction), 5 (student learning conversations), and 6 (student active engaged learning) and tested the strength of relationship between the sum of percentages and WESTEST 2 proficiency rates for the four core content areas. A similar grouping of IPI categories was used by Collins and Valentine in their study of successful and unsuccessful middle schools. In their study they found that students in highly successful schools, based on student achievement data, are "significantly more likely to be engaged in higher order thinking with teachers who are actively teaching the students" (Valentine, 2005, p. 13).

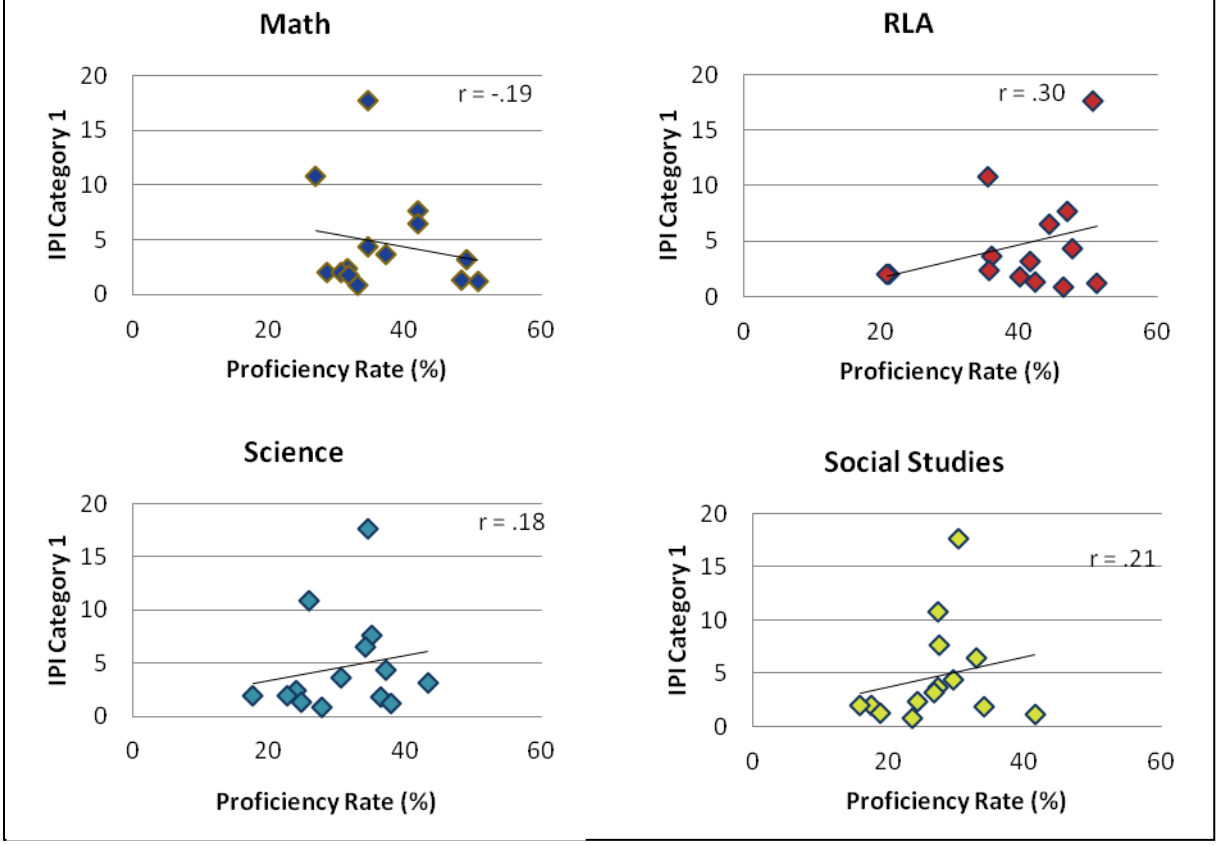
Correlation analysis for SIG schools (Figure 10) showed a strong positive relationship between the sum of percentages for IPI Categories 4, 5, and 6 and mathematics as well as reading/language arts proficiency rates (r = .52 and .59, respectively). Although the correlation between the three IPI categories and science proficiency rates was not as strong, results indicate that it was still on the very high end of the moderate positive strength level (r = .49). The strength of relationship between the three categories and proficiency rates for social studies was found to be the weakest out of the four content areas, however, results indicate there is a moderate positive relationship (r = .34). All core content areas except social studies proficiency were statistically significant at the 0.05 level (Table 15, page 36). Category 4 (teacher-led instruction) by itself showed a strong positive relationship with reading/ language arts proficiency rate (r = .58), a moderate positive relationship with science proficiency rates (r = .30), and a weak relationship with proficiency rates for mathematics and social studies (r = .21 and r = .19, respectively). Only the proficiency rate for reading/ language arts was statistically significant at the 0.05 level (Table 15, page 36).

Figure 10. Scatter Plots of 2010-2011 Proficiency Rates and IPI Categories 4, 5, & 6



As previously mentioned, analysis of IPI data by Collins and Valentine (Valentine, 2011) indicates a strong inverse relationship between complete disengagement (Category 1) and student achievement. Obviously, this is the level of engagement that authors strongly recommend schools attempt to eliminate from their instructional practices. Although Collins and Valentine limited their analysis to the content areas of communication arts and mathematics, we extended our analysis to also include proficiency rates for science and social studies (Figure 11). The results of the analyses show only a weak negative relationship between disengagement and proficiency rates in mathematics (r = -.19). Contrary to Collins and Valentine's findings, results of correlation analyses between IPI Category 1 data and proficiency rates for the other three core content areas, reading/language arts, science, and social studies, show a weak to a low-moderate positive relationship (r = .30, .18, and .21 respectively). None of the correlations were statistically significant at the 0.05 level (Table 15, page 36).

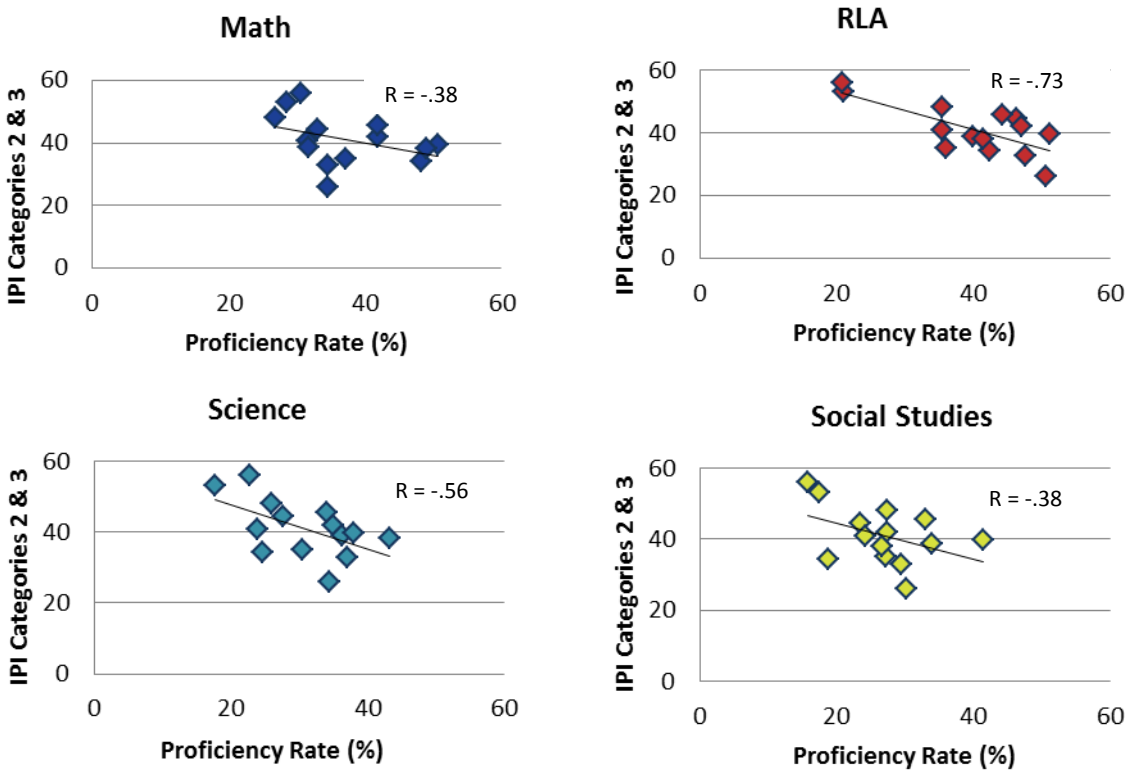
Figure 11. Scatter plots of 2010-2011 Proficiency Rates and IPI Category 1



In addition to attempting to eliminate complete student disengagement (Category 1), Collins and Valentine (Valentine, 2011) encourage schools to reduce to a minimum the percentages in two other IPI coding categories; student work with teacher not engaged (Category 2) and student work with teacher engaged (Category 3). Both categories indicate students are doing seatwork and higher-order learning is not evident.

We tested the strength of relationship between the sum of percentages of IPI Categories 2 (student work with teacher not engaged) and 3 (student work with teacher engaged) and proficiency rates on WESTEST 2 for the four core content areas (Figure 12). The results indicate that there is an inverse relationship between the two categories combined and proficiency rates for all content areas covered by WESTEST 2. Categories 2 and 3 combined show a strong negative relationship with reading/language arts and science proficiency rates (r = -.73, -.56 respectively) and a moderate negative correlation with mathematics and social studies proficiency rates (r = -.38, -.38 respectively). The correlation was only statistically significant at the 0.01 level for reading/language arts but at the 0.05 level for science. Correlation analysis conducted to test the strength of relationship between each category and proficiency rates indicated a moderate to strong negative relationship for Category 3 with reading/language arts, science, and social studies proficiency rates (r = -.74, -.52, and -.42, respectively). The correlation between reading/language arts proficiency rates and Category 3 was significant at the 0.01 level but only significant at the 0.05 level for science proficiency rates (Table 15, page 36). On the other hand, Category 2 showed a strong negative relationship with mathematics proficiency rates (r = -.61) and the result was statistically significant at the 0.05 level (Table 15, page 36).

Figure 12. Scatter plots of 2010-2011 Proficiency Rates and IPI Categories 2 & 3



Comparison of baseline IPI data-SIG schools and typical baseline IPI data

Collins and Valentine's (Valentine, 2009) *typical data* are based on IPI profiles from schools that have participated in projects undertaken by the Middle Level Leadership Cen-

ter, and data voluntarily submitted to the Center by schools using the IPI process. It is worth repeating here the caveat mentioned previously: Typical IPI profile data are derived by averaging data submitted by schools of various enrollment sizes, programmatic levels, and locales (i.e., rural, suburban, and urban) that were *not* selected systematically. Therefore, *typical IPI profiles* cannot be considered to be statistically representative of a population and conclusions should not be drawn from these data. However, data presented below provide us with the opportunity to compare baseline IPI data for SIG schools with available baseline data for other schools using the IPI process.

Baseline *core IPI data*—that is, IPI data for core subject areas—for SIG schools and a baseline core IPI profile for a typical school are displayed in Table 16, Table 17, and Table 19. Table 16 presents data for each SIG school along with the average profile for the group, while Table 17, Table 18, and Table 19 present the data by programmatic levels. Data presented in Table 18 is the *total IPI data*—that is, IPI data for *all* subject areas—for middle schools rather than core IPI data (an explanation is provided below with the discussion for Table 18). All four of these tables include data for each IPI category as well as four groupings of IPI categories, three of which are discussed in the section above. The only grouping of IPI categories presented in the tables that has not been discussed is the grouping of Categories 1, 2, and 3. We did not find it necessary to report the results separately since the sum of percentages of these three categories is technically the inverse of the sum of percentages of Categories 4, 5, and 6, which are already addressed in the section above.

Overall, the comparison shows that average baseline data for IPI categories are roughly similar between SIG schools and the typical school profile (Table 16). Data for the two groups in three of the six individual categories (i.e., Categories 1, 4, and 5) were within 1.5 percentage points. The table does, however, show some differences. It is interesting to note that two of the three IPI categories in which the two groups showed greater variation were Categories 6 and 3, both of which showed some level of relationship with WESTEST 2 proficiency rates for SIG schools (previously discussed). For IPI Category 6, which showed a positive relationship with proficiency rates for mathematics, science, and social studies, the average baseline data were approximately 3 percentage points lower for SIG schools. For Category 3, which showed a negative relationship with proficiency rates for reading/language arts, science, and social studies, the average baseline data were more than 6 percentage points higher.

Student engagement in higher order/deeper learning (IPI Categories 5 and 6) is lower in SIG schools by more than 3 percentage points, while data for IPI Categories 2 and 3 are higher by approximately 3.5 percentage points. The sum of percentages of Categories 4, 5, and 6 is also lower by more than 4 percentage points in SIG schools. The latter grouping of IPI categories (Categories 4, 5, and 6) has shown a moderate to strong positive relationship with proficiency rates for all four content areas on WESTEST 2 (see section above). In four of the 14 SIG schools, the sum of percentages of Categories 1, 2, and 3 were greater than the sum of percentages of Categories 4, 5, and 6 (Table 16). This indicates that continued support and encouragement is needed to increase student engagement in higher order/deeper learning and teacher-led instruction for all SIG schools and particularly for the four schools in which the percentages of IPI categories 4, 5, and 6 are lower.

Table 16. Comparison of Baseline IPI Data for All SIG Schools and Typical School Profile

	Percentages for IPI categories and sums of percentages for category groupings (core content areas)									
School	6	5	4	3	2	1	5-6	2-3	4-5-6	1-2-3
Average baseline data (typical)*	10.83	4.76	43.35	28.35	9.05	3.66	15.59	37.40	58.94	41.06
Average baseline data (SIG Schools)	7.95	4.29	42.01	34.66	6.35	4.67	12.24	41.01	54.25	45.68
School I	7.56	**1.68**	47.06	19.33	6.72	**17.65**	**9.24**	26.05	56.3	43.7
School N	**2.45**	**3.68**	55.21	31.90	3.07	3.68	**6.13**	34.97	61.34	38.65
School B	**0.81**	**2.42**	51.61	33.06	**11.29**	0.81	**3.23**	**44.35**	54.84	45.16
School E	15.20	6.40	35.20	33.60	7.20	2.40	21.60	40.80	56.80	43.20
School K	9.21	9.21	46.05	31.58	2.63	1.32	18.42	34.21	64.47	35.53
School O	**4.64**	**3.61**	32.99	**37.11**	**10.82**	**10.82**	**8.25**	**47.93**	**41.24**	**58.75**
School H	**0.76**	**0.76**	48.85	**37.40**	4.58	7.63	**1.52**	**41.98**	50.37	49.61
School C	5.69	**3.25**	39.02	**39.84**	5.69	6.50	**8.94**	**45.53**	**47.96**	**52.03**
School L	13.94	**4.24**	41.21	32.12	6.67	1.82	18.18	38.79	59.39	40.61
School J	13.02	5.33	40.83	35.5	4.14	1.18	18.35	39.64	59.18	40.82
School M	10.00	**3.00**	31.00	**50.00**	3.00	2.00	13.00	**53.00**	**44.00**	**55.00**
School A	19.05	4.76	34.92	34.92	3.17	3.17	23.81	38.09	58.73	41.26
School D	6.01	7.65	49.18	21.86	**10.93**	4.37	13.66	32.79	62.84	37.16
School F	**3.00**	**4.00**	35.00	**47.00**	9.00	2.00	**7.00**	**56.00**	**42.00**	**58.00**

*Data collected from 264 schools based on 40,314 observations conducted between November 1996 and March 2009 (Valentine, 2009).
Red indicates areas of concern.

Table 17 presents core baseline IPI data for elementary schools. More caution is necessary due to the small number of SIG schools at this programmatic level (n = 4) and conclusions should not be drawn from these data. The comparison shows that average baseline data for the sum of percentages of IPI Categories 4, 5, and 6 is almost 5 percentage points lower in SIG schools (56.8%) compared to typical school profile (61.7%). However, it should be noted here that three of the four SIG elementary schools have comparable numbers to the average typical baseline data for the three categories, an average of 61.7%. The overall percentage is lower due mainly to a very low percentage for the three IPI categories for School F (42.0%). Limiting the comparison to student engagement in higher order/ deeper learning (IPI Categories 5 and 6), it is evident that, overall, the percentages of the two categories combined are lower by over 4.5 percentage points in SIG elementary schools (12.5% verses 17.1%). Again, a closer look at individual SIG elementary schools shows us that two of the four schools, School K and School J actually have higher sum of percentages than the typical average baseline data (18.4% and 18.4%, respectively).

Individually, the greatest variation is found in Categories 2, 3, and 6. Although a lower percentage in teacher disengagement (Category 2) is a positive finding for SIG schools (4.7%) compared with the typical elementary school (7.1%), percentage of Category 3 was higher in SIG schools (36.5%) by a large margin compared to typical elementary schools

(28.4%). Moreover, percentage of Category 6 was lower in SIG schools by almost 6 percentage points compared to the typical elementary school (6.9% compared to 12.8%).

Table 17. Comparison of Baseline IPI Data for SIG Elementary Schools and Typical Elementary School Profile

	Percentages for IPI categories and sums of percentages for category groupings (core content areas)									
Elementary school	**6**	5	4	3	2	1	5-6	2-3	4-5-6	1-2-3
Average baseline data (typical)*	12.78	4.36	44.53	28.39	7.05	2.89	17.14	35.44	61.67	39.33
Average baseline data (SIG Schools)	6.92	5.56	44.27	36.50	4.71	2.05	12.48	41.21	56.75	43.25
School N	2.45	3.68	55.21	31.90	3.07	3.68	6.13	34.97	61.34	38.65
School K	9.21	9.21	46.05	31.58	2.63	1.32	18.42	34.21	64.47	35.53
School J	13.02	5.33	40.83	35.50	4.14	1.18	18.35	39.64	59.18	40.82
School F	3.00	4.00	35.00	47.00	9.00	2.00	7.00	56.00	42.00	58.00

*Data collected from 94 schools between September 2004 and February 2009 (Valentine, 2009).

For the middle school level, using total (all content areas) rather than core (mathematics, reading/language arts, science, and social studies) IPI data in their analysis, Collins and Valentine provide profiles for highly successful and very unsuccessful schools (Valentine, 2005). Their data came from six highly successful middle schools that were part of a national study and five unsuccessful schools that had achievement data in the bottom 5% of an unnamed mid-Western state. Unlike the typical IPI data discussed in this section, data for the two groups of schools were collected under controlled research conditions and have been statistically tested for significance (statistically significant differences at the level of .05 between the two groups of schools are italicized in the table below). Baseline data for SIG middle schools (n = 8) and the typical profile presented in Table 18 is therefore also for total IPI data. This allowed researchers to compare IPI profiles of SIG schools not only with the typical profile but also with IPI profiles of schools with histories of high and low achievement, despite their small sample size.

The test for differences for the following IPI categories was statistically higher in high achieving schools: (a) total higher-order/deeper learning (IPI Categories 5 and 6); (b) total sum of percentages of Categories 4, 5, and 6; and (c) Category 5. On the other hand, the presence of Categories 3 and 1 individually and Categories 1, 2, and 3 in combination were significantly higher in schools with histories of low achievement. The sum of percentages of Categories 1 and 2 was also higher in unsuccessful schools (22.0% versus 9.5%) and the difference was significant at the level of .05. Table 18 shows that the presence of Category 6 is higher in successful schools but the difference tested significant only at the .10 level. Presence of teacher-led instruction, Categories 3 and 4 combined, showed little variation between successful schools (57.8%), and unsuccessful schools (61.6%).

Overall, the comparison between SIG schools and the typical school profile shows that average baseline data for total IPI categories are similar, within 3 percentage points (Table 18). The only exception is in the presence of total higher-order/deeper learning which is lower in SIG schools by a little more than 3 percentage points (16.5% versus 13.4%). It is

worth pointing out that the presence of total higher-order/deeper learning was also lower in SIG schools (13.4%) compared to very unsuccessful schools (16.2%). This was mainly due to two outlier schools (School B and School H) that have considerably lower numbers for this category. Excluding the values for these two schools, the average for the remaining six SIG middle schools (16.7%) is very similar to the percentages found for both typical school profile (16.51) and unsuccessful schools (16.2).

One interesting observation, also highlighted by Valentine (2005), is that in highly successful schools the ratio of the sum of percentages for Categories 4, 5, and 6 to Categories 1, 2, and 3 is approximately 3:1, whereas in unsuccessful school the ratio is almost 1:1. The ratio for SIG schools and the typical school profile is almost identical and both fare only slightly better than unsuccessful schools (Table 18). Looking at the data a bit more closely, however, it is evident that some SIG middle schools perform considerably better than unsuccessful schools and to some extent their SIG peers. The ratio of the sum of percentages for Categories 4, 5, and 6 to Categories 1, 2, and 3 for School I, School L, and School D are almost exactly 3:2 or greater.

Table 18. Comparison of Baseline IPI data for SIG Middle Schools With Typical, Very Successful, and Very Unsuccessful Middle School Profiles

	Percentages for IPI categories and sums of percentages for category groupings (total content areas)									
Middle school	6	5	4	3	2	1	5-6	2-3	4-5-6	1-2-3
Average baseline data (typical)*	11.45	5.06	39.71	30.71	9.13	3.93	16.51	39.85	56.22	43.78
Average baseline data (SIG Schools)	9.65	3.73	42.39	32.30	6.96	4.98	13.38	39.25	55.77	44.23
Very Successful**	29.3	*3.3*	40.5	*17.3*	8.5	*1.0*	*32.6*	25.8	*73.10*	*26.80*
Very Unsuccessful***	16.0	*0.2*	33.2	*28.4*	13.6	*8.4*	*16.2*	42.0	*49.40*	*50.40*
School I	9.38	3.13	46.88	20.00	6.88	13.75	12.51	26.88	59.39	40.63
School B	1.50	3.01	51.13	33.08	10.53	0.75	4.51	43.61	55.64	44.36
School E	14.67	5.33	33.33	35.33	8.67	2.67	20.00	44.00	53.33	46.67
School H	1.48	0.74	48.15	37.04	4.44	8.15	2.22	41.48	50.37	49.63
School C	8.45	2.82	38.03	38.73	5.63	6.34	11.27	44.36	*49.30*	*50.70*
School L	13.94	4.24	41.21	32.12	6.67	1.82	18.18	38.79	59.39	40.61
School A	17.86	3.57	32.14	41.67	2.38	2.38	21.43	44.05	53.57	46.43
School D	9.95	6.97	48.26	20.40	10.45	3.98	16.92	30.85	65.18	34.83

*Data collected from 44 schools between September 2004 and February 2009 (Valentine, 2005)

**Based on student achievement data from National Association of Secondary Schools Principals' National Study collected in 2002 from six schools in six states (Valentine, 2005)

***Based on student achievement data from Project ASSIST collected in five schools in 2003 in a mid-Western state (Valentine, 2005)

Since the number of SIG high schools in our sample is very small (n = 2) rather than discussing the average baseline data for the two schools, researchers compared the schools to each other and to the typical average baseline data. The sum of percentages of Categories 4, 5, and 6 is very similar for the two SIG schools but this statistic for both schools was more than 10 percentage points lower than the typical school profile (Table 19). On the other

hand, the sum of percentages of Categories 2 and 3 combined was higher in both SIG schools compared to the typical school profile, although School M was more than 5 percentage points higher than School O. The sums of percentages of higher-order/deeper learning (Categories 5 and 6) were almost identical for School M and the typical school profile but much lower in School O. It should be noted that a lower percentage of Category 6 in School O accounts for the difference between the two SIG schools for this grouping of categories (4.6% versus 10.0%).

Table 19. Comparison of Baseline IPI data for SIG High Schools and Typical High School profile

	Percentages for IPI categories and sums of percentages for category groupings (core content areas)									
High School	6	5	4	3	2	1	5-6	2-3	4-5-6	1-2-3
Average baseline data (typical)*	8.23	4.88	41.81	27.84	11.93	5.31	13.11	39.77	54.92	45.08
Average baseline data (SIG Schools)	7.32	3.31	32.00	43.56	6.91	6.41	10.63	50.47	42.62	56.88
School O	4.64	3.61	32.99	37.11	10.82	10.82	8.25	47.93	41.24	58.75
School M	10.00	3.00	31.00	50.00	3.00	2.00	13.00	53.00	44.00	55.00

*Data collected from 64 schools between September 2004 and February 2009 (Valentine, 2009)

Looking at the broader IPI category of disengagement (Category 1 and 2), the presence of disengagement, both student and teacher, is considerably more pronounced in School O, where it is more than 4 percentage points higher than the typical school profile and more than four times higher compared to School M The presence of Category 3 was much higher in the two SIG schools compared to the typical school profile and it is higher by more than 12 percentage points in School M compared to School O.

Effective School Practice Survey Results

As described earlier (see Methods section, page 10), the Effective School Practice Survey measures perceptions about the fidelity of schools' operations to tenets of effective school practice in five domains: (a) Leadership and Decision-Making; (b) Professional Development; (c) Parents and Learning; (d) Curriculum, Assessment, and Instructional Planning; and (e) Classroom Instruction. For the majority of items, the 4-point Likert-type response scale was as follows: 1 = *Rarely, if ever*, 2 = *Sometimes*, 3 = *Often*, and 4 = *Consistently*. Two additional categories, *I don't know* and *Too early to tell*, were also included but not used to calculate mean scores for the items. Nine survey items used 4-point Likert-type response scales that were item specific. A higher mean score (3.00 or above) indicates that the school is implementing effective school practice(s) with high fidelity (i.e., 1 = *little or no fidelity*, 2 = *limited fidelity*, 3 = *above average fidelity*, and 4 = *strong fidelity)*.

Overall average rating by domain

On a scale of 1 to 4, the overall average rating for all schools across all five domains was 2.68, (Table 20), a rating that falls between *limited* and *above average* fidelity. Mean domain scores for all schools ranged between 2.33 (Parents and Learning) and 2.88 (Leadership and Decision Making). There was little variation among the domains (SD = 0.22). In addition to Parents and Learning, one more domain (Curriculum, Assessment, and Instruc-

tional Planning) had a mean score very near the overall average (M = 2.70). Three out of the five domains had average scores that were above the overall mean (Classroom Instruction, Professional Development, and Leadership and Decision Making). Yet, none of the domain mean scores reached 3.0 on the 4-point scale which would have indicated strong fidelity.

Table 20. Mean Scores by Effective School Practice Survey Domain

	Leadership and decision making	Professional development	Parents and learning	Curriculum, assessment, and instructional planning	Classroom instruction	Overall mean* (SD)
Mean	2.88	2.77	2.31	2.70	2.74	2.68 (0.22)

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity

Figure 13 displays overall mean scores for each school, while Table 21 provides scores for each domain and overall mean scores for each school. The mean score for the five domains among schools ranged between 1.89 for School M and 3.65 for School A (SD = 0.51), with 54.5% of the schools (n = 6) scoring below the group average (M = 2.68) and only one school (School M) scoring below 2.00. It should be noted, however, only two schools (School A and School J) scored above the threshold for above average fidelity. The greatest amount of variation was found in the Parents and Learning domain, with the lowest score at 1.33 and highest score at 3.67 (SD = 0.76).

Figure 13. Mean Effective School Practice Survey Scores by School

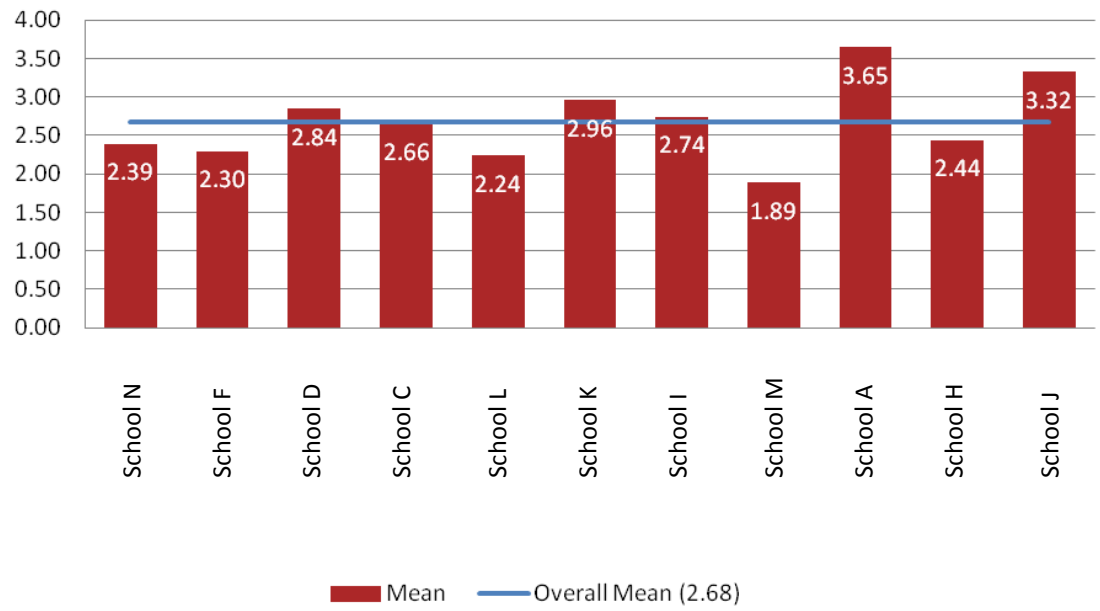


Table 21. Average Effective School Practice Survey Scores by Domain by School

School (number of raters)	Leader-ship and decision making	Professional development	Parents and learning	Curriculum, assessment, and instructional planning	Classroom instruction	Mean* (SD)
School N (1)	3.07	2.00	2.00	2.40	2.46	2.39 (0.44)
School F (2)	2.68	3.17	1.33	1.90	2.40	2.30 (0.71)
School D (3)	3.36	3.22	2.42	2.58	2.64	2.84 (0.42)
School C (4)	3.16	2.92	1.83	2.82	2.55	2.66 (0.51)
School L (1)	2.31	2.00	2.00	2.50	2.38	2.24 (0.23)
School K (2)	3.00	3.00	3.00	2.92	2.87	2.96 (0.06)
School I (3)	3.25	2.61	2.50	2.63	2.71	2.74 (0.29)
School M (2)	2.14	2.00	1.33	2.00	1.97	1.89 (0.32)
School A (2)	3.68	3.50	3.67	3.75	3.63	3.65 (0.09)
School H (1)	2.22	2.00	ND**	3.00	2.54	2.44 (0.43)
School J (1)	2.85	4.00	3.00	3.20	3.57	3.32 (0.47)
Mean	2.88	2.77	2.31	2.70	2.74	2.68
(SD)	(0.50)	(0.70)	(0.76)	(0.53)	(0.50)	(0.51)

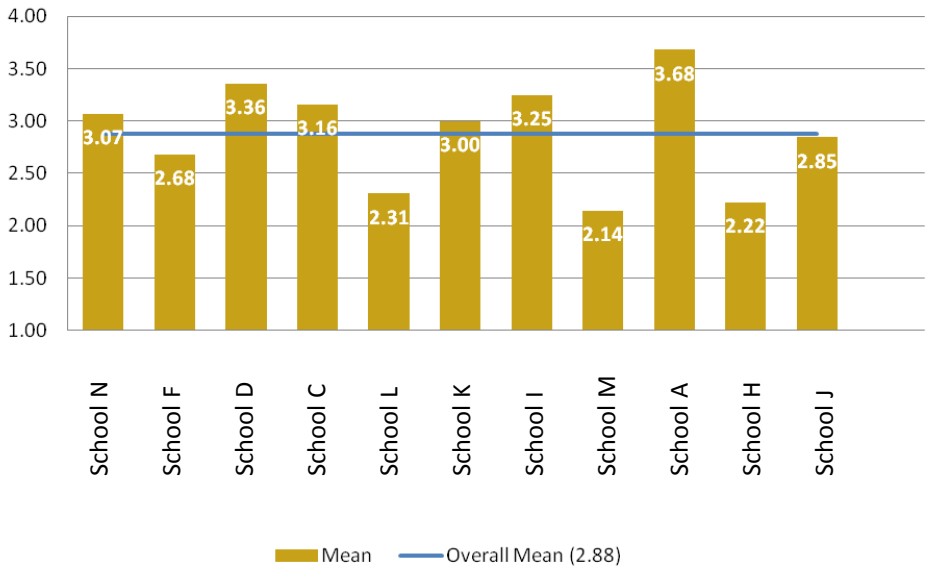
*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity
**ND = No data

It also should be noted that 59 out of 132 possible responses (44.7%) from all raters to items in the Parents and Learning domain were "I don't know" (Table B- 2, page 125). SEA Title I coordinators accounted for the majority of these responses, with 73.6% of their responses falling under this response category. On the other hand, there were no "I don't know" responses from school improvement specialists, and only 16.7% of LEA Title I directors responded in that way.

Leadership and decision making

The average score for the 11 schools for the domain *Leadership and Decision Making* was 2.88 (Figure 14; Table B- 3, page 126) which was the highest average score among the five domains. Scores ranged between 2.14 (School M) and 3.68 (School A) with a standard deviation of 0.50, which is the smallest standard deviation among average item scores of the five domains. Six schools (54.5%) scored at or above the threshold of above average fidelity (3.0). The average score for each of the 14 items for the domain ranged between 1.99 and 3.80. Out of the 14 items, 57.1% (n = 8) received scores lower than 3.0. The item measuring the frequency with which the school leadership team communicates DP21 data with teams and teachers received the lowest score, while the item that asked how often collaborative teams meet received the highest score. Overall, results indicate this is an area of strength for SIG schools as a group. However, there are a few schools with scores that are relatively low for this domain.

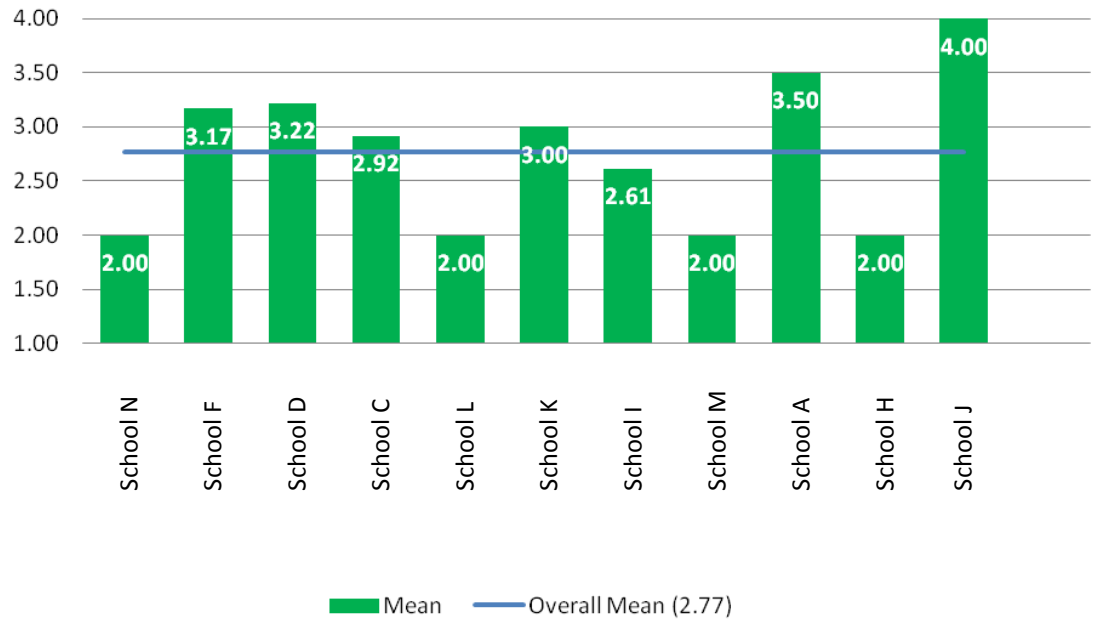
Figure 14. Average Effective School Practice Survey Scores for Leadership and Decision Making by School



Professional development

For the domain *Professional Development*, the average domain score for all schools was 2.77, which was the second highest average score among the five domains (Figure 15; Table B- 4, page 127). Average scores ranged between 2.00 (School N, School L, School M, and School H) and 4.00 (School J) with a standard deviation of 0.70, which was the second largest standard deviation among average scores for the five domains. Very little variation was found among average scores for the three items/indicators for the professional development domain, 2.71, 2.74, and 2.98, respectively, and all indicators fall below the threshold of above average fidelity. Similar to the dimension discussed above, overall, this appears to be an area of some strength when one looks at the average score for the domain in comparison to the other domains. However, only five schools (45.5%) have scores that are above 3.0, indicating above average fidelity (Figure 15). Furthermore, four schools have a score of 2.00, which indicates limited fidelity in this domain of effective school practice.

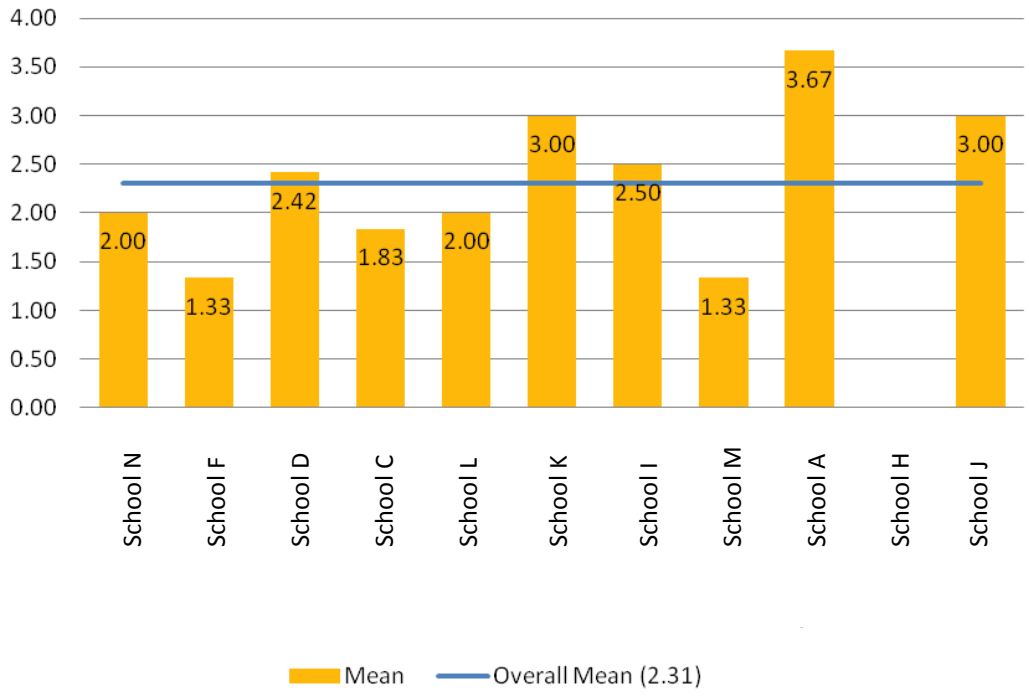
Figure 15. Average Effective School Practice Survey Scores for Professional Development by School



Parents and Learning

The average score for the 11 schools for the domain *Parents and Learning* was 2.31 (Figure 16; Table B- 5, page 128) and was the lowest average score among the five domains (this may be due to the lack of sufficient data for some and/or all indicators for four out of the 11 schools). Average scores ranged between 1.33 (School F and School M) and 3.67 (School A) with a standard deviation of 0.76, which is the highest standard deviation among average scores of the five domains. Only three out of the 10 schools (30.0%) for which data were available scored above the threshold of above average fidelity. There were no data for School H (Figure 16). Average score for each of the six items/indicators for this domain fell below 3.0 ranging between 1.81 and 2.63. The item measuring the frequency with which parents are provided opportunities to meet with each other to share their child-rearing concerns and successes received the lowest score (1.81), while the item that asked how often parents receive jargon-free communication about learning standards, their children's progress, and the parents' role in their children's school success received the highest score (2.63). Based on the available data this area appears to be an area of particular concern.

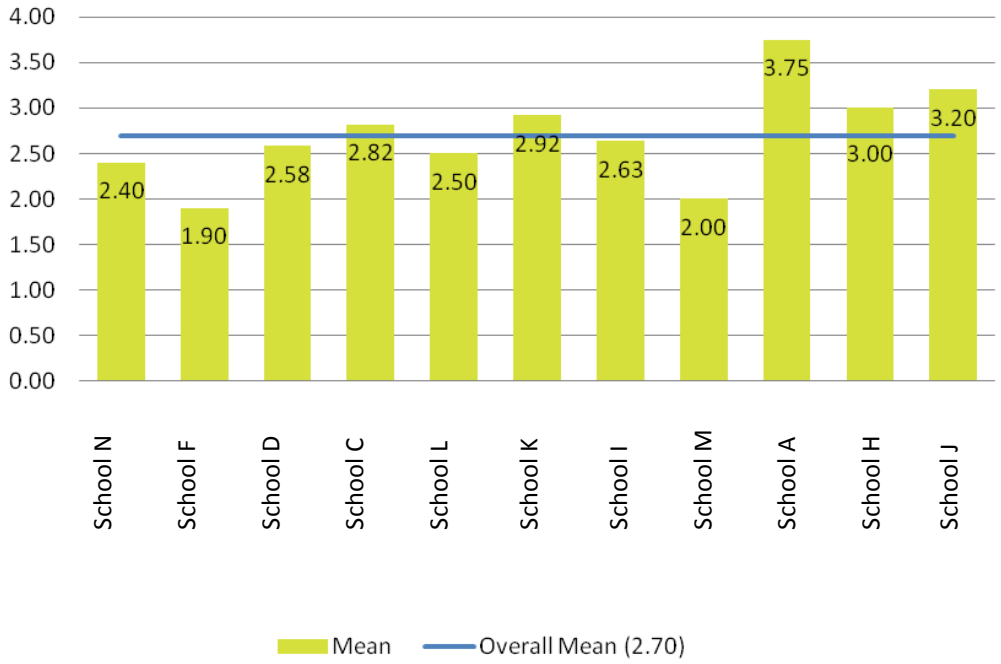
Figure 16. Average Effective School Practice Survey Scores for Parents and Learning by School



Curriculum, Assessment, and Instructional Planning

Mean scores for the domain *Curriculum, Assessment, and Instructional Planning* ranged between 1.90 (School F) and 3.75 (School A), with an average score of 2.70 for all schools and a standard deviation of 0.53 (Figure 17, Table B- 6, page 129). The average score for this domain is the second lowest average score among the five domains. Only three out of 11 schools (27.3%) had mean scores at or above the threshold of above average fidelity. Eight out of 10 indicators have scores below 3.0 with average scores ranging between 2.19 for the frequency with which WVEIS data is validated after initial entry and 3.27 for the frequency with which units of instruction include standards-based objectives and criteria for mastery. Data indicates this is an area of some concern for SIG schools.

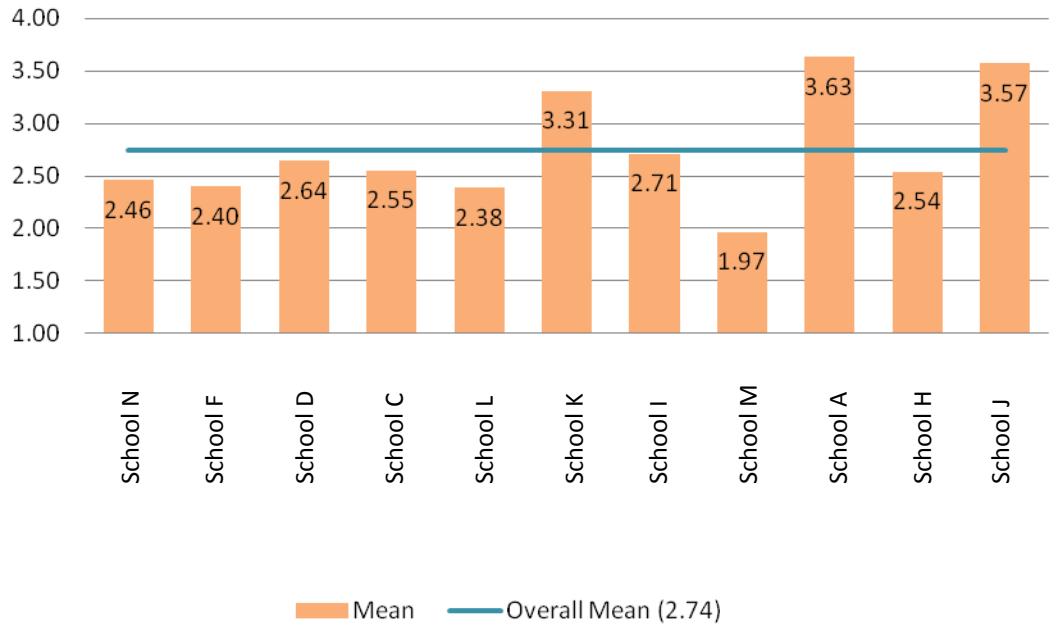
Figure 17. Average Effective School Practice Survey Scores for Curriculum, Assessment, and Instructional Planning, by School



Classroom Instruction

The mean score for the *Classroom Instruction* domain was 2.74 (Figure 18, Table B-7, page 130), which was the third lowest average score among the five domains. Average scores among the 11 schools ranged between 1.97 (School M) and 3.63 (School A) with a 0.53 SD. Only three of the 11 schools (27.5%) had mean scores above the threshold for above average fidelity. Twelve out of 15 indicators (80%) had scores below 3.0. Mean scores for the 15 indicators ranged between 2.36 for the frequency with which educators *encourage students to paraphrase, summarize, and relate,* and 3.47 for the frequency with which educators *develop weekly lesson plans based on aligned units of instruction with clearly stated lesson topics, themes, and learning objectives*. Results indicate this is an area of some concern.

Figure 18. Average Effective School Practice Survey Scores for Classroom Instruction, by School



Overall Mean Score by Programmatic Level

Out of the 11 schools for which data were available, 45.5% were elementary schools (n = 5), 45.5% were middle schools (n = 5) and [redacted] (Table 22). The mean scores for all five domains among the three programmatic level ranged between 1.89 (high school) and 2.76 (elementary and middle school), with a mean score of 2.45 (SD = 0.50). All three programmatic levels scored below 3.0, the threshold for above average fidelity. It should be pointed out here that there was only one high school in the sample. There was no difference between elementary and middle schools.

Figure 19. Overall Effective School Practice Survey Mean Scores, by Programmatic Level

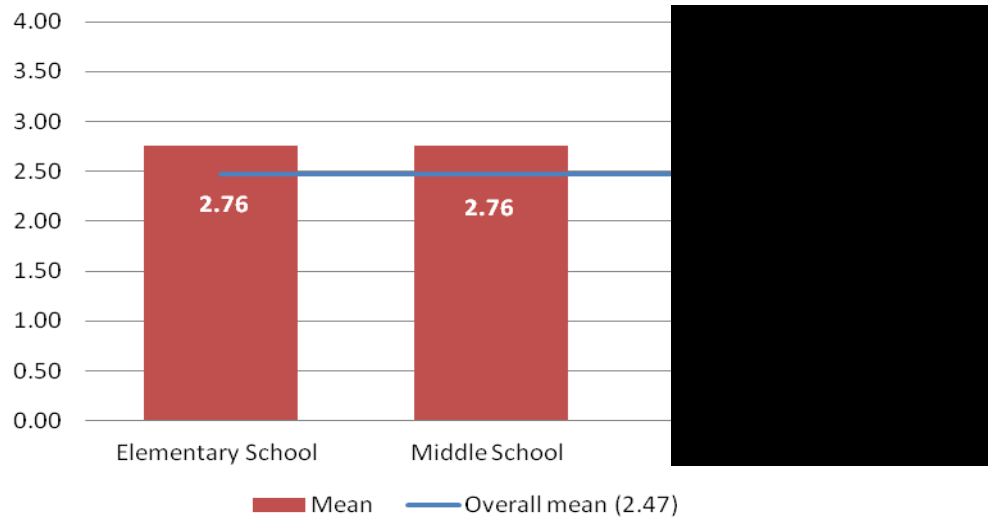


Table 22. Average Effective School Practice Survey Domain Scores, by Programmatic Level

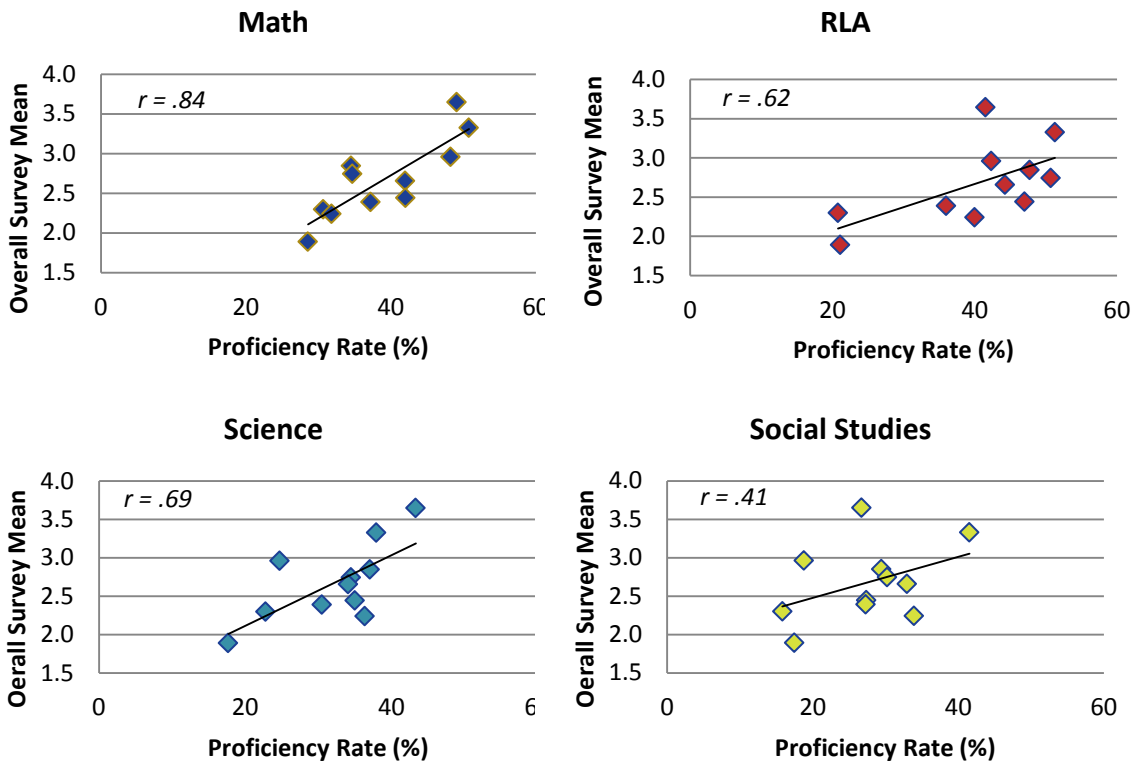
School (Number of schools)	Leadership and decision making (14 Items)	Professional development (3 Items)	Parents and learning (6 Items)	Curriculum, assessment, and instructional planning (10 Items)	Classroom instruction	Mean* (SD)
Elementary School (5)	2.99	3.08	2.35	2.34	2.79	2.76 (0.30)
Middle School (5)	2.92	2.61	2.50	2.94	2.76	2.76 (0.19)
[redacted]						
Mean	2.69	2.56	2.06	2.51	2.51	2.45
SD	0.47	0.54	0.63	0.48	0.47	0.50

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity

Correlation results between Effective School Practice Survey and achievement data

Results showed that mean scores were positively correlated with proficiency rates for all content areas at least at the high end of a moderate relationship (Figure 20). Correlation with social studies proficiency rates was the weakest (r = .41) and was not statistically significant at the 0.05 level (Table 23). Proficiency rates for science and reading/language arts were strongly correlated, at .69 and .62 respectively (Table 23). Correlation for reading/language arts was statistically significant at the 0.05 level, whereas the correlation for Science was significant at the 0.01 level. The strongest relationship was found between mathematics proficiency rates and survey mean scores (r = .84). This correlation was significant at the 0.01 level.

Figure 20. Scatter plots of 2010-2011 Proficiency Rates and Effective School Practice Survey Mean Score



We also tested the strength of relationship between each of the five domains that make up the Effective School Practice Survey and proficiency rates for the four content areas. Results indicate that three domains, Parents and Learning, Curriculum, Assessment, and Instructional Planning, and Classroom Instruction, were positively correlated with all content areas at least at the moderate level. With the exception of the correlation between social studies proficiency rates and the three domains, all other correlations were statistically significant at least at the 0.05 level (Table 23). Results indicate that the domain with the strongest correlation with proficiency rates overall is Curriculum, Assessment, and Instruc-

tional Planning, which tested statistically significant at the 0.01 level with proficiency rates for the three content areas.

Table 23. Bivariate Correlation Between Proficiency Rates and Effective School Practice Survey Domains

	Content area proficiency rate correlations			
	Mathematics	Reading/ language arts	Science	Social studies
Mean	.836 **	.617 *	.687 **	.407
Leadership and Decision making	.436	.426	.520	.195
Professional Development	.592 *	.306	.359	.267
Parents and Learning	.832 **	.683 *	.677 *	.362
Curriculum, Assessment, and Instructional Planning	.882 **	.691 **	.768 **	.484
Classroom Instruction	.898 **	.554 *	.559 *	.349

*Correlation is significant at the 0.05 level (1-tailed).
**Correlation is significant at the 0.01 level (1-tailed).

The two remaining domains, Leadership and Decision Making and Professional Development, were positively correlated with proficiency rates for three content areas (mathematics, reading/language arts, and science) at least at the moderate level (Table 23). However, only the correlation between Professional Development and mathematics proficiency rates were statistically significant.

Teacher attendance and instructional minutes

To examine the effect of the SIG program on the quantity of instruction over the course of the program, we will document trends for the rate of teacher attendance and number of instructional minutes per school year. Data from the 2010-2011 school year will serve as baseline data in subsequent reports.

Evaluation Question 3

EQ3. To what extent has the SIG program resulted in improved academic achievement among students in participating schools?

To address EQ3, we conducted two analyses—one using student-level assessment data, which examined effects of the SIG program on academic achievement statewide; and another conducted at the school level, which provided feedback to each individual SIG school. Both analyses were conducted using the same sample of SIG and non-SIG comparison schools and WESTEST 2 scores for mathematics and reading/language arts.

Grade-level cohort findings

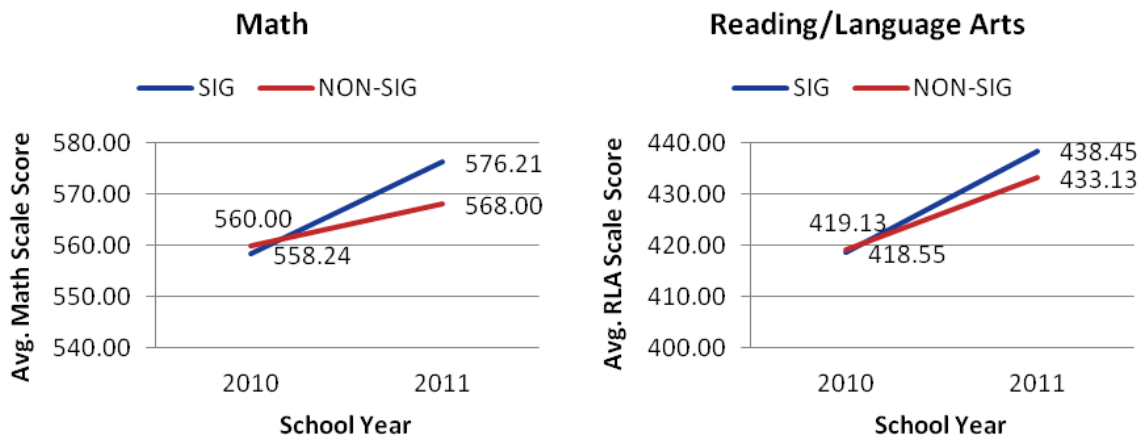
For each grade level cohort, we conducted a series of RM ANOVAs to test for significant group*time interactions during the implementation year. As noted in the Methods section (page 18), we did not conduct RM ANOVAs when our matching process was unsuccessful, as evidenced from *t* test results. When significant group*time interactions were present, we determined which group of students experienced higher gains over time. We also analyzed the data to determine if there were any main effects for time and group in-

dependently, though this was not pertinent to the research question. Results are presented below by grade level. Detailed model statistics, including observed statistical power and effect size for each grade level cohort and content area appear in Table C- 1 through Table C-16, in Appendix C (page 131).

Grade 3–4 cohort

For Grade 3–4 mathematics, the interaction between time and group was significant, $F(1, 619) = 10.686$, $p = .001$, as was the effect for time, $F(1, 619) = 61.645$, $p < .000$. However, the main effect for group was not statistically significant, $F(1, 619) = .791$, $p < .374$. *These results indicate that there is a significant difference in third to fourth grade mathematics gains between students in SIG and non-SIG schools during the first implementation year. The students in SIG schools realized greater gains from Grade 3–4 during the implementation year, than students in non-SIG schools.* Figure 21 illustrates that students in SIG schools gained approximately 18 scale score points, whereas students in non-SIG comparison schools gained approximately 8 scale score points.

Figure 21. Grade 3–4 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011



For Grade 3–4 reading/language arts, the interaction between time and group was also significant, $F(1, 619) = 5.345$, $p = .021$, as was the effect for time, $F(1, 619) = 177.053$, $p = .000$. However, the main effect for group was not statistically significant, $F(1, 619) = .556$, $p < .456$. *These results indicate that there is a significant difference in third to fourth grade reading/language arts gains between students in SIG and non-SIG schools during the first implementation year. The students in SIG schools realized greater gains from Grade 3–4 during the implementation year, than students in non-SIG schools.* Figure 21 illustrates that students in SIG schools gained approximately 20 scale score points, whereas students in non-SIG comparison schools gained approximately 15 scale score points.

Grade 4–5 cohort

For Grade 4–5 mathematics, the interaction between time and group was not significant, $F(1, 492) = .029$, $p = .866$, nor was the effect for group, $F(1, 492) = 1.03$, $p = .312$. However, the main effect for time was statistically significant, $F(1, 492) = 130.20$, $p < .000$. *These results indicate that there is not a significant difference in fourth to fifth grade mathematics gains between students in SIG and non-SIG schools during the first implementation year*. Figure 22 illustrates that students in SIG schools gained approximately 22 scale score points, whereas students in non-SIG comparison schools gained approximately 23 scale score points.

Figure 22. Grade 4–5 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011

For Grade 4–5 reading/language arts, the interaction between time and group was not significant, $F(1, 492) = .254$, $p = .614$, nor was the effect for group, $F(1, 492) = .597$, $p = .440$. However, the main effect for time was statistically significant, $F(1, 492) = 112.868$, $p < .000$. *These results indicate that there is not a significant difference in fourth to fifth grade reading/language arts gains between students in SIG and non-SIG schools during the first implementation year*. Figure 22 illustrates that students in SIG schools gained approximately 16 scale score points, whereas students in non-SIG comparison schools gained approximately 15 scale score points.

Grade 5–6 cohort

For Grade 5–6 mathematics, the interaction between time and group was not significant, $F(1, 377) = 1.817$, $p = .179$, nor was the effect for group, $F(1, 377) = .640$, $p = .424$. However, the main effect for time was statistically significant, $F(1, 377) = 24.902$, $p < .000$. *These results indicate that there is not a significant difference in fifth to sixth grade mathematics gains between students in SIG and non-SIG schools during the first implementation year***.** Figure 23 illustrates that students in SIG schools gained approximately 10 scale score points, whereas students in non-SIG comparison schools gained approximately 18 scale score points.

For Grade 5–6 reading/language arts, the interaction between time and group was significant, $F(1,378) = 5.260$, $p = .022$, as was the effect for time, $F(1, 378) = 36.232$, $p < .000$. However, the main effect for group was not statistically significant, $F(1, 378) = .724$, $p = .395$. *These results indicate that there is a significant difference in fifth to sixth grade reading/language arts gains between students in SIG and non-SIG schools during the first implementation year. The students in SIG schools realized greater gains from Grade 5–6 during the implementation year, than students in non-SIG schools.*

Figure 23. Grade 5–6 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011

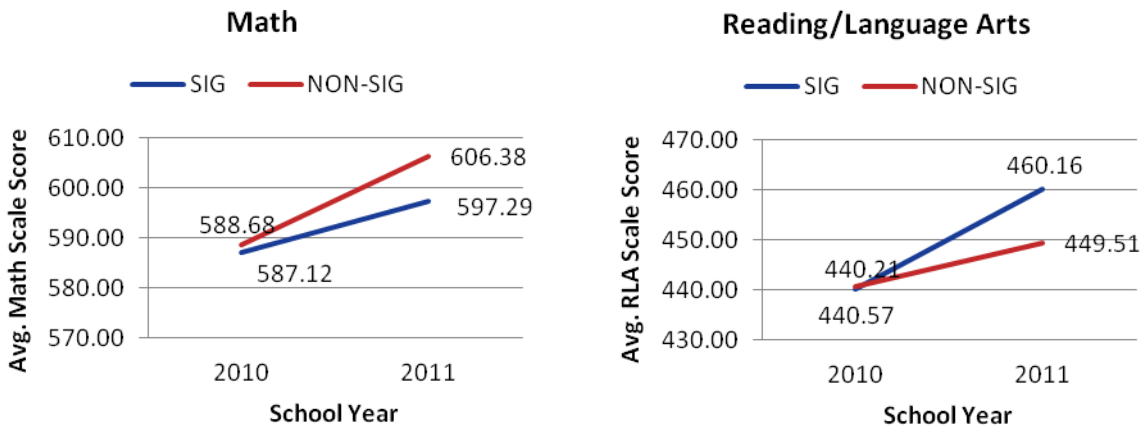
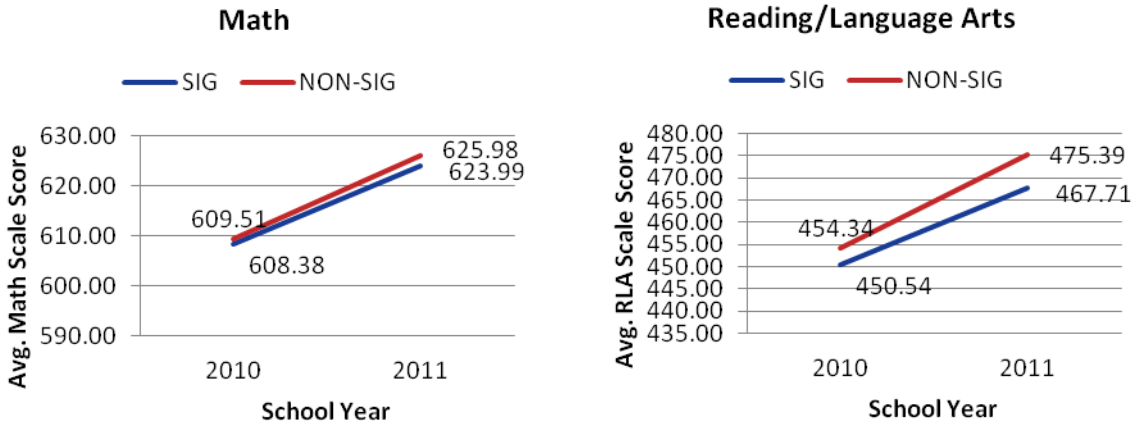


Figure 23 illustrates that students in SIG schools gained approximately 20 scale score points, whereas students in non-SIG comparison schools gained approximately 9 scale score points.

Grade 6–7 cohort

For Grade 6–7 mathematics, the interaction between time and group was not significant, $F(1, 1474) = .181$, $p = .670$, nor was the effect for group, $F(1, 1474) = .492$, $p = .483$. However, the main effect for time was statistically significant, $F(1, 1474) = 256.059$, $p < .000$. *These results indicate that there is not a significant difference in fifth to sixth grade mathematics gains between students in SIG and non-SIG schools during the first implementation year***.** Figure 24 illustrates that students in SIG schools gained approximately 16 scale score points, whereas students in non-SIG comparison schools gained approximately 16 scale score points.

Figure 24. Grade 6–7 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011

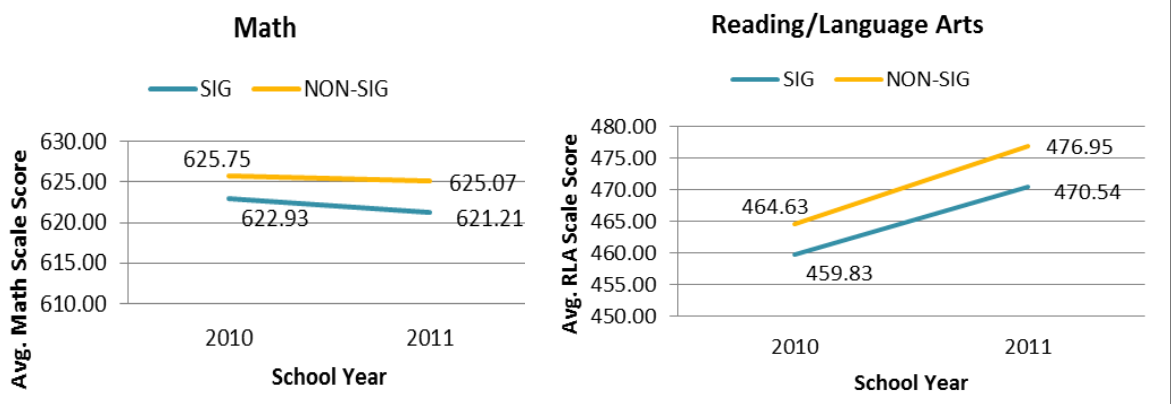


For Grade 6–7 reading/language arts, the interaction between time and group was significant, $F(1, 1475) = 5.258$, $p = .022$, as was the effect for group, $F(1, 1475) = 7.316$, $p = .007$, and time , $F(1, 1475) = 506.845$, $p < .000$. *These results indicate that there is a significant difference in fifth to sixth grade reading/language arts gains between students in SIG and non-SIG schools during the first implementation year. The students in non-SIG comparison schools realized greater gains from Grade 6–7 during the implementation year, than students in SIG schools.* Figure 24 illustrates that students in SIG schools gained approximately 17 scale score points, whereas students in non-SIG comparison schools gained approximately 21 scale score points.

Grade 7–8 cohort

For Grade 7–8 mathematics, the interaction between time and group was not significant, $F(1, 1426) = .192$, $p = .662$, nor was the effect for group, $F(1, 1426) = 1.643$, $p = .20$ or time $F(1, 1426) = 1.042$, $p = .308$. *These results indicate that there is not a significant difference in seventh to eighth grade mathematics gains between students in SIG and non-SIG schools during the first implementation year.* Figure 25 illustrates that students in SIG schools lost approximately 2 scale score points, whereas students in non-SIG comparison schools lost approximately 1 scale score point.

Figure 25. Grade 7–8 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011

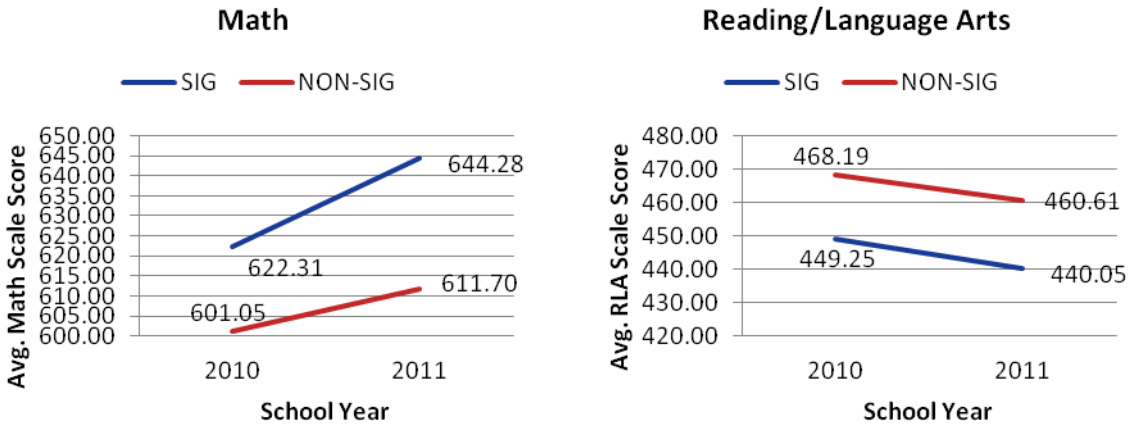


For Grade 7–8 reading/language arts, our two groups were not adequately matched at baseline. Therefore, we did not meet the assumption necessary to trust the results of a subsequent RM ANOVA. We provide descriptive data here only as a point of reference. We caution against drawing any inferences about the effectiveness of the SIG program from these results. Figure 25 illustrates that students in SIG schools gained approximately 11 scale score points, whereas students in non-SIG comparison schools gained approximately 12 scale score points.

Grade 8–9 cohort

For Grade 8–9 mathematics, our two groups were not adequately matched at baseline. Therefore, we did not meet the assumption necessary to trust the results of a subsequent RM ANOVA. We provide descriptive data here only as a point of reference and caution against drawing any inferences about the effectiveness of the SIG program from these results. Figure 26 illustrates that students in SIG schools gained approximately 22 scale score points, whereas students in non-SIG comparison schools gained approximately 11 scale score points.

Figure 26. Grade 8–9 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011

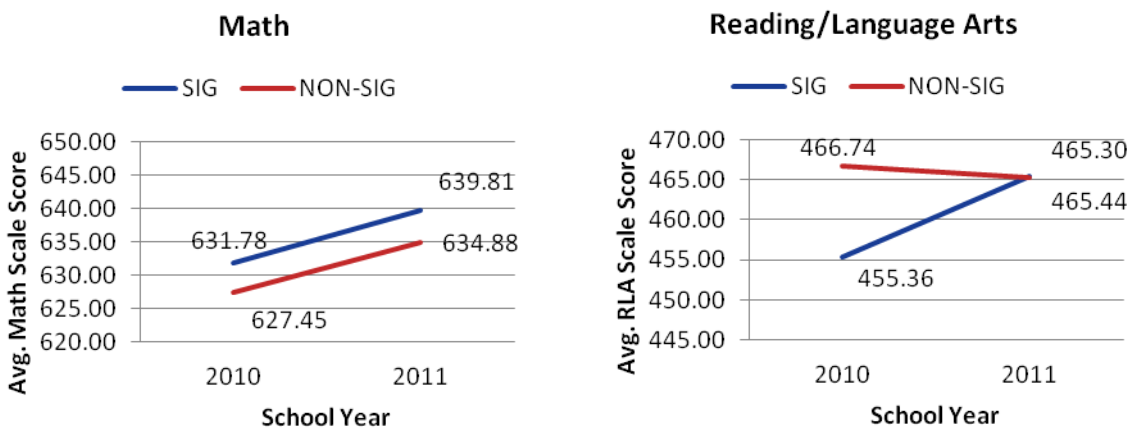


For Grade 8–9 reading/language arts, our two groups were not adequately matched at baseline. Therefore, we did not meet the assumption necessary to trust the results of a subsequent RM ANOVA. We provide descriptive data here only as a point of reference, and caution against drawing any inferences about the effectiveness of the SIG program from these results. Figure 26 illustrates that students in SIG schools lost approximately 9 scale score points, whereas students in non-SIG comparison schools lost approximately 7 scale score points.

Grade 9–10 cohort

For Grade 9–10 mathematics, the interaction between time and group was not significant, $F(1, 670) = .019$, $p = .891$, nor was the effect for group, $F(1, 670) = 1.741$, $p = .187$. However, the effect for time was statistically significant, $F(1, 670) = 12.629$, $p < .000$. *These results indicate that there is not a significant difference in ninth to 10th grade mathematics gains between students in SIG and non-SIG schools during the first implementation year.* Figure 27 illustrates that students in SIG schools gained approximately 8 scale score points, whereas students in non-SIG comparison schools gained approximately 7 scale score points.

Figure 27. Grade 9–10 SIG versus Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011



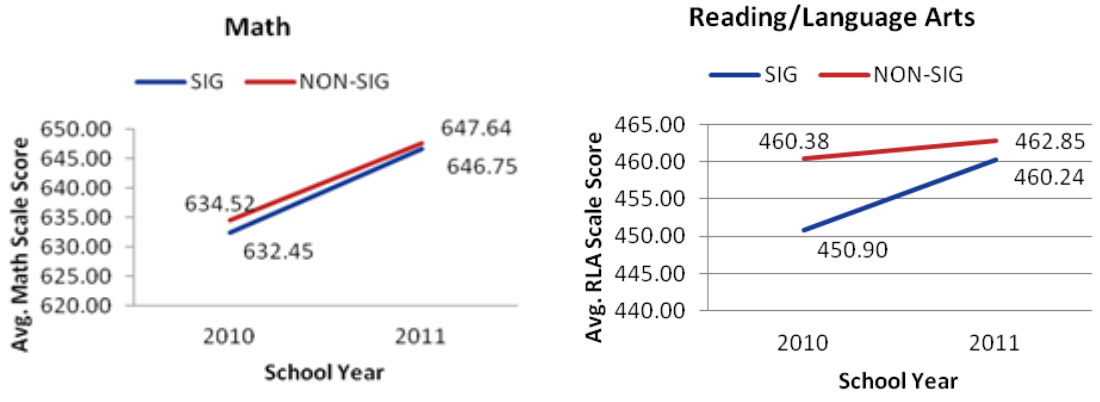
For Grade 9–10 reading/language arts, our two groups were not adequately matched at baseline. Therefore, we did not meet the assumption necessary to trust the results of a subsequent RM ANOVA. We provide descriptive data here only as a point of reference and caution against drawing any inferences about the effectiveness of the SIG program from these results. Figure 27 illustrates that students in SIG schools gained approximately 10 scale score points, whereas students in non-SIG comparison schools lost approximately 1 scale score point.

Grade 10–11 cohort

For Grade 10–11 mathematics, the interaction between time and group was not significant, $F(1, 575) = .088$, $p = .767$, nor was the effect for group, $F(1, 575) = .185$, $p = .667$. However, the effect for time was statistically significant, $F(1, 575) = 48.021$, $p < .000$. *These results indicate that there is not a significant difference in 10th to 11th grade mathematics gains between students in SIG and non-SIG schools during the first implementation year***.** Figure 28 illustrates that students in SIG schools gained approximately 14 scale score points, whereas students in non-SIG comparison schools gained approximately 13 scale score points.

For Grade 10–11 reading/language arts, the interaction between time and group was not significant, $F(1, 578) = 2.752$, $p = .098$, nor was the effect for group, $F(1, 578) = 1.455$, $p = .228$. However, the effect for time was statistically significant, $F(1, 578) = 8.143$, $p = .004$. *These results indicate that there is not a significant difference in 10th to 11th grade mathematics gains between students in SIG and non-SIG schools during the implementation year*. Figure 28 illustrates that students in SIG schools gained approximately 9 scale score points, whereas students in non-SIG comparison schools gained approximately 2 scale score points.

Figure 28. Grade 10–11 SIG vs. Non-SIG Gains for Mathematics and Reading/Language Arts, 2010–2011



School-Level findings

Description of participants

We began our analysis of school-level results by compiling a data file containing all students tested at SIG and non-SIG comparison schools during SY2009, SY2010, and SY2011. The file contained 27,833 records. Figure 29 indicates that the number of students increased consistently between school years with the fewest student records available in SY2009 (n = 8,987) and most in SY2011 (n = 9,493).

We also examined the distribution of students by grade level. The dispersion appears to have been stable across school years with Grades 3 through 5 comprising approximately 20% of the sample each year, Grades 6 through 8 comprising between 55% and 57%, and Grades 9 through 11 comprising between 23% and 24%. Figure 30 provides a visual representation.

Figure 29. Number of Students in SIG and Non-SIG Comparison Schools by School Year

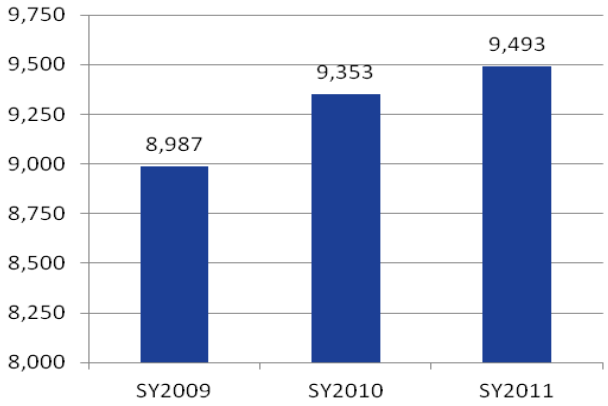
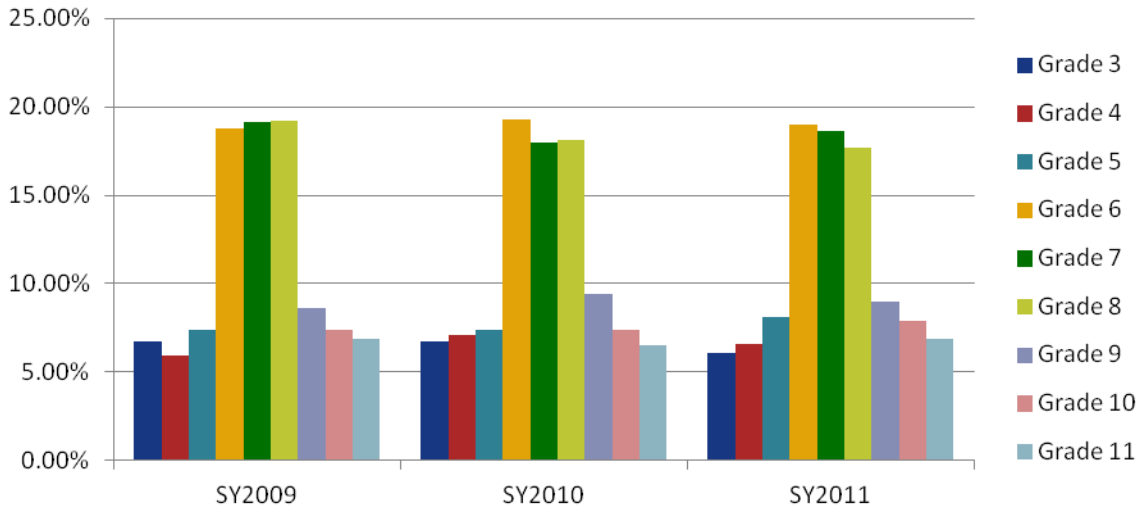


Figure 30. Distribution of Students in SIG and Non-SIG Comparison Schools by School Year



We also reviewed the composition of our sample by school year based upon free and reduced price lunch eligibility, special education eligibility, and gender. Figure 31, Figure 32, and Figure 33 provide overviews of our findings. For all three variables, there was stability across school years.

Figure 31. Free and Reduced Price Lunch Eligibility in SIG and non-SIG Comparison Schools by School Year

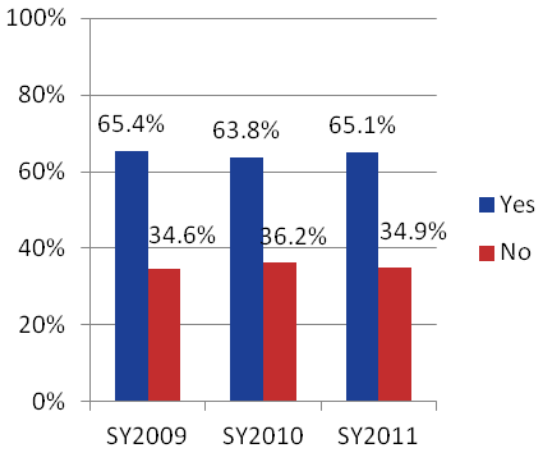


Figure 32. Special Education Eligibility in SIG and non-SIG Comparison Schools by School Year

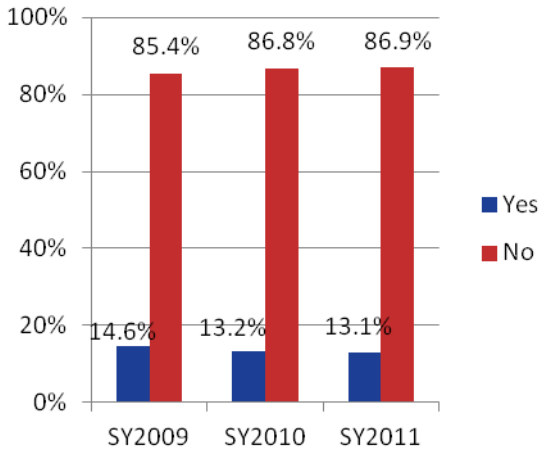
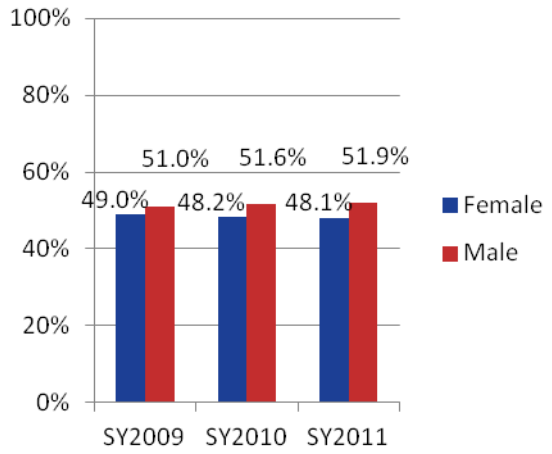


Figure 33. Gender in SIG and non-SIG Comparison Schools by School Year



Determining the school-wide percent proficient

We used information about each individual student's proficiency level on each subtest (i.e., 1 = *Novice*, 2 = *Partial Mastery*, 3 = *Mastery*, 4 = *Above Mastery*, 5 = *Distinguished*) to develop a binary proficiency status variable for that student for both reading/language arts and mathematics (i.e., 1 = proficient, 0 = not proficient). As is the convention in West Virginia for calculating adequate yearly progress (AYP), we considered all students at or above the Mastery cut score to be proficient, and all those below the cut score to be not proficient.

It should be noted that WV developed revised cut scores for proficiency in 2010. For our achievement profiles, we applied the revised and currently operational 2010 cut scores to the 2009 testing data. This allowed us to examine proficiency trends using a common metric. Furthermore, we should also note that *the proficiency rates published here do not always correspond directly with those reported for NCLB accountability. This is because, for our analyses, we calculated school level proficiency rates using all available assessment data for students who were tested, not just data for those students for whom the school is formally held accountable.* Using this information, we then calculated overall proficiency rates for each school. We did so for all students across all grade levels during SY2009, SY2010, and SY2011. These proficiency rates represent the proportion of students in a given school across all grade levels who scored at or above the cut score for proficiency during that school year. We also calculated corresponding proficiency rates for any NCLB-reported subgroups that exceeded 50 students during any of the 3 years for which WESTEST 2 data were available[9] for a given school. We then plotted this information for 14 pairs consisting of a Cohort 1 SIG and a non-SIG comparison school (n = 14). We provide a brief narrative of the gains that have occurred over time in each school.

Results for individual schools

We examined trends in proficiency rates in reading/language arts and mathematics for two time periods (1) the period between SY2009 and SY2011 and (2) the period encompassing the first year of SIG implementation—from SY2010–SY2011.

We first summarized program level trends in proficiency rate changes for the group of Cohort 1 SIG schools and the group of non-SIG comparison schools. *We found that all Cohort 1 SIG schools (100%) have experienced net proficiency gains in reading/language*

[9] If a school had 50 or more students in a given subgroup 1 year across all grades, but fewer than 50 in that same subgroup during another year, we presented the trend data for all 3 years as long as the subgroup size did not drop so low in any year as to become unreliable (fewer than 25 students).

arts and mathematics between SY2009 and SY2011. The amount of improvement in these schools ranged between +7.60% and +26.16% in reading/language arts and between +3.47 and +24.23% in mathematics. The median increase for Cohort 1 SIG schools during this period was +14.51% in reading/language arts and +10.64% in mathematics.

By way of comparison, we also examined trends for the 14 comparison group schools. Twelve (86%) also made net gains during the 3-year period in reading/language arts while 11 schools (79%) made net gains in mathematics. The magnitude of the changes ranged from -9.27% and +37.92% in reading/language arts and between -9.35% and +30.86% in mathematics. The median increase was +8.39% in reading/language arts and +6.97% in mathematics.

Table 24 provides summary data for this 3-year period. These data, while compelling, are descriptive in nature and we should note that, for this analysis, we followed these schools across the entire period of time that WESTEST 2 has been administered, including a non-SIG implementation period (SY2009–SY2010) so the entirety of these net gains cannot possibly be attributed to the SIG program.

Table 24. Improvement in Proficiency SY2009–SY2010 by Group

Group	# (%) of schools improving SY2009–SY2011	Median improvement SY2009–SY2011	Range of performance SY2009–SY2011
Reading/Language Arts			
SIG Cohort 1	14 (100%)	+14.51%	+7.60% to +26.16%
Non-SIG comparison	12 (80%)	+8.39%	-9.27% to +37.92%
Mathematics			
SIG Cohort 1	14 (100%)	+10.64%	+3.47 to +24.23%
Non-SIG comparison	11 (78%)	+6.97%	-9.35 to +30.86%

We were also quite interested in examining gains during the first implementation year of the SIG program (i.e., between SY2010 and SY2011). For this time period, we found that 12 of the 14 Cohort 1 SIG schools (86%) experienced gains in reading/language arts proficiency rates. The median increase for the group was +8.62%, but schools ranged from a -1.08% to +14.82%. During the same period, 11 of the 14 Cohort 1 SIG schools (79%) experienced gains in mathematics proficiency rates. The median increase was +5.39% but, again schools ranged in their performance from -2.70% to +19.12%.

By way of comparison, we also examined proficiency rates for the 14 comparison schools. Twelve (86%) also experienced gains in reading/language arts proficiency rates despite not being involved in the SIG program. The median reading/language arts increase for this group was +8.28% with a range of -8.84 to +17.58%. However, only seven of the comparison schools (50%) experienced gains in mathematics during the same period. The median improvement was less than +1% (+.11%) for this group, and the range was -11.66% to +5.17%.

Table 25 provides summary data for the first SIG implementation year. Again, these data, while compelling, are descriptive in nature. Furthermore, they do not fully encompass the SIG implementation period which continues until the conclusion of SY2012.

Table 25. Improvement in Proficiency SY2010–SY2011 by Group

Group	Number (percent) of schools improving SY2010–SY2011	Median Improvement SY2010–SY2011	Range of Performance SY2010–SY2011
Reading/Language Arts			
SIG Cohort 1	12 (86%)	+8.62%	-1.08% to +14.82%
Non-SIG comparison	12 (86%)	+8.28%	-8.84 to +17.58%
Mathematics			
SIG Cohort 1	11 (79%)	+5.39%	-2.70 to +19.12%
Non-SIG comparison	7 (50%)	+.11%	-11.66 to +5.17%

Profiles of individual Cohort 1 SIG schools and their associated comparison schools appear in Figure 34 through Figure 47 in the following pages. At this time, the profiles are completely descriptive in nature. We do not intend for schools to use information from these preliminary achievement profiles to make any definitive conclusions about the success of their SIG program nor to infer that the differences between their achievement trends and those of their comparison school are statistically significant. Our intention in providing the profiles is only to provide a point of contrast for the types of gains occurring in non-SIG schools that are neither enacting the reforms required by the grant nor receiving the large investments of resources provided by the program.

School I

Figure 34 indicates that School I has experienced continuous proficiency gains for all students in both reading/language arts (+21.72%) and mathematics (+10.10%) between SY2009 and SY2011. There was a large increase of 14.82% in reading/language arts and 5.55% of the increase in mathematics during the SIG implementation year—between SY2010 and SY2011. However, as a point of contrast, School P has also experienced an overall improvement in both content areas (+8.31% in reading/language arts and +7.10% in mathematics) and experienced gains of similar magnitude during the SIG implementation year, despite not being a SIG school. See Table 26 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Table D- 1.

Figure 34. Achievement Profile for School I: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

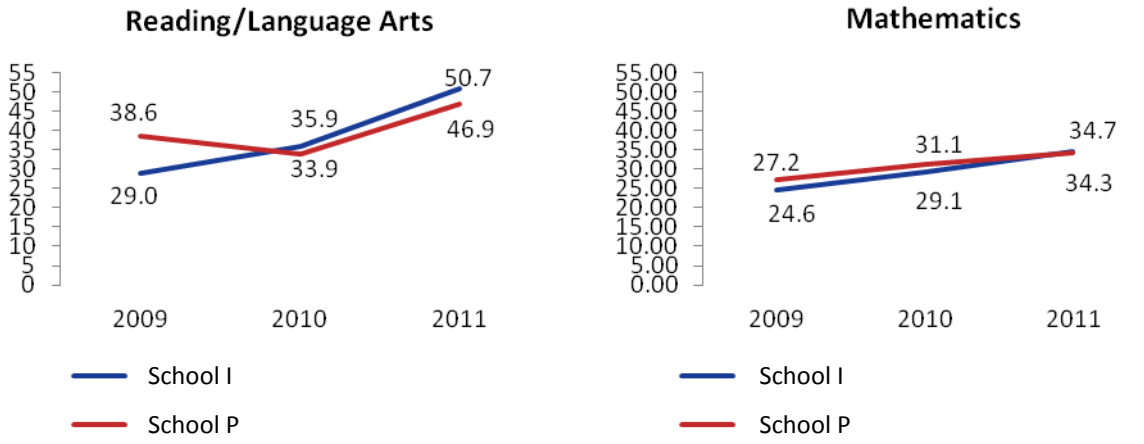


Table 26. Reading/Language Arts and Mathematics Gains for All Students in School I and School P

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School I	+6.90	+14.82	+4.55	+5.55
School P	-4.72	+13.03	+3.94	+3.16
School Δ	+11.63	+1.79	+0.61	+2.39

School N

Figure 35 indicates that School N has experienced continuous proficiency gains for all students in mathematics (+11.18%) and achieved a net gain in reading/language arts (+11.12%) between SY2009 and SY2011. There was a marginal decline in reading/language arts proficiency (-1.08%) and a gain of 7.78% in mathematics during the initial SIG implementation year. As a point of contrast, School Q has experienced an overall decline of -9.27% in reading/language arts and a marginal gain of +1.90% in mathematics between SY2009 and SY2011. The school has declined in proficiency in both content areas between SY2010 and SY2011. See Table 27 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 35. Achievement Profile for School N: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

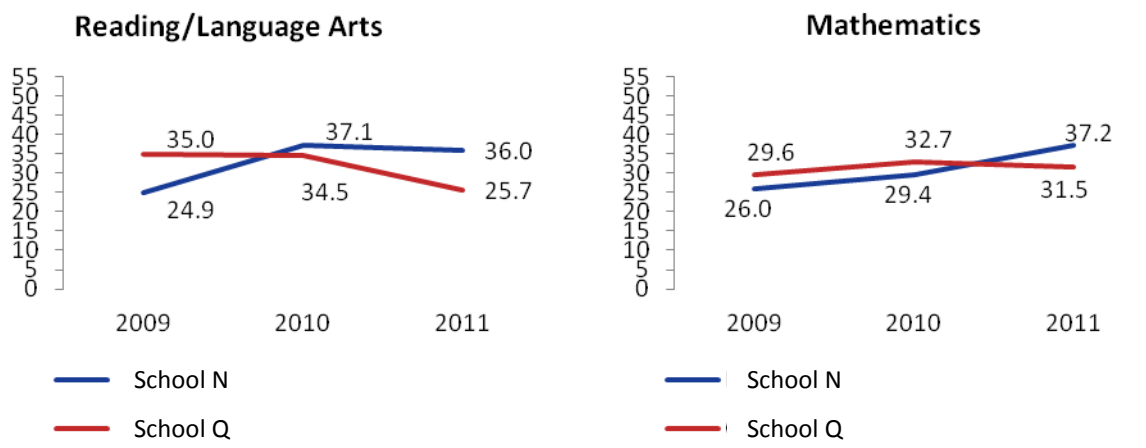


Table 27. Reading/Language Arts and Mathematics Gains for All Students in School N and School Q

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School N	+12.20	-1.08	+3.40	+7.78
School Q	-0.43	-8.84	+3.17	-1.27
School Δ	+12.63	+7.76	+0.23	+9.06

School K

Figure 36 indicates that School K has experienced continuous proficiency gains for all students in both reading/language arts (+17.60%) and mathematics (+21.50%) between SY2009 and SY2011. There was an increase of 8.18% in reading/language arts and a gain of 19.12% in mathematics during the initial SIG implementation year. As a point of contrast, School R has experienced a comparable overall increase of 15.94% in reading/language arts and a smaller gain of 6.87% in mathematics between SY2009 and SY2011. The school has increased comparably in reading/language arts proficiency (+6.13%), but has not improved as dramatically as School K in mathematics (+1.80%) between SY2010 and SY2011. See Table 28 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 36. Achievement Profile for School K: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

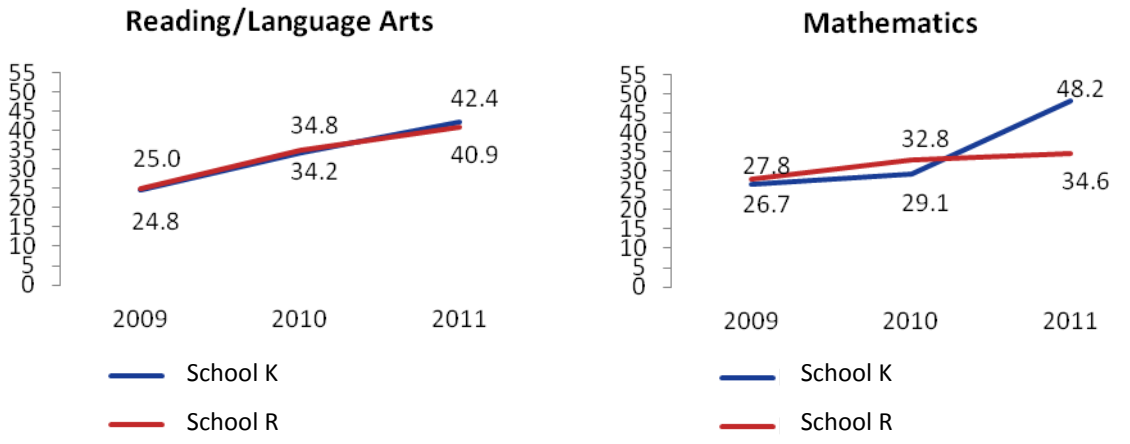


Table 28. Reading/Language Arts and Mathematics Gains for All Students in School K and School R

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School K	+9.42	+8.18	+2.38	+19.12
School R	+9.81	+6.13	+5.07	+1.80
School Δ	-0.39	+2.05	-2.69	+17.32

School B

Figure 37 indicates that School B has experienced continuous proficiency gains for all students in both reading/language arts (+20.16%) and mathematics (+10.08%) between SY2009 and SY2011. There was an increase of 8.54% in reading/language arts and a gain of 5.69% in mathematics during the initial SIG implementation year. As a point of contrast, School S has experienced an overall decrease of -1.93% in reading/language arts and a decrease of -7.35% in mathematics between SY2009 and SY2011. The school has increased marginally in reading/language arts proficiency (+3.64%), and has decreased in mathematics (-5.40%) between SY2010 and SY2011. See Table 29 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 37. Achievement Profile for School B: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

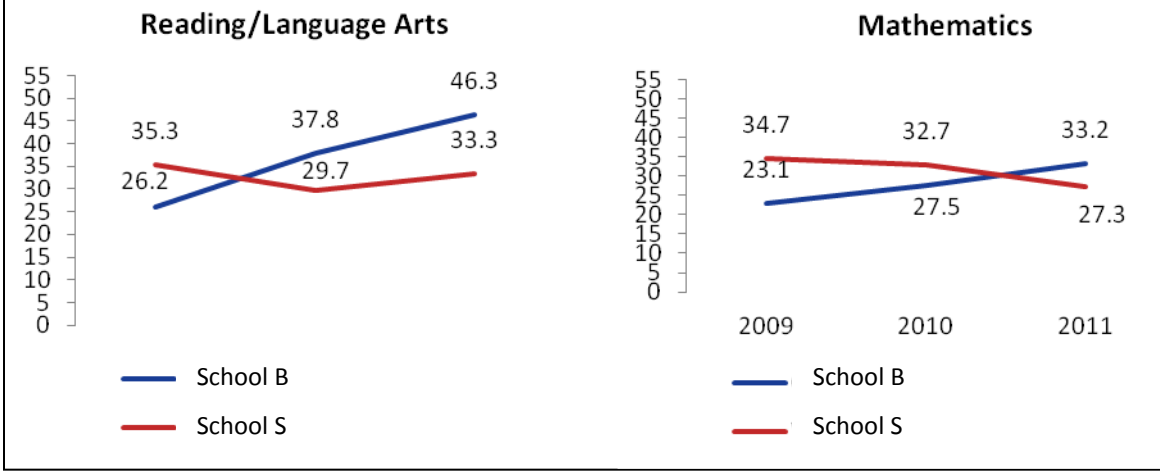


Table 29. Reading/Language Arts and Mathematics Gains for All Students in School B and School S

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009–SY2010	SY2010–SY2011	SY2009–SY2010	SY2010–SY2011
School B	+11.62	+8.54	+4.40	+5.69
School S	-5.56	+3.64	-1.95	-5.40
School Δ	+17.19	+4.90	+6.35	+11.08

School E

Figure 38 indicates that School E has experienced continuous proficiency gains for all students in reading/language arts (+15.63%) and an overall gain in mathematics (+9.78%) between SY2009 and SY2011. There was an increase of 10.13% in reading/ language arts and a marginal decrease of -.88% in mathematics during the initial SIG implementation year. As a point of contrast, School T has experienced an overall increase of 8.47% in reading/language arts and 4.41% in mathematics between SY2009 and SY2011. The school has increased in reading/language arts proficiency (+6.87%), and has decreased in mathematics (-4.45%) between SY2010 and SY2011. See Table 30 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 38. Achievement Profile for School E: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

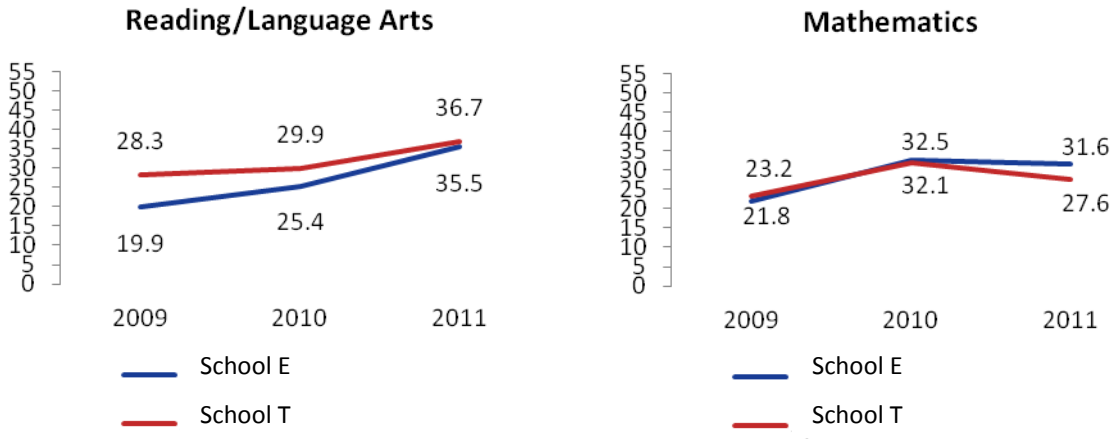


Table 30. Reading/Language Arts and Mathematics Gains for All Students in School E and School T

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009–SY2010	SY2010–SY2011	SY2009–SY2010	SY2010–SY2011
School E	5.50	10.13	10.67	-0.88
School T	1.60	6.87	8.87	-4.45
School Δ	3.90	3.25	1.80	3.57

School H

Figure 39 indicates that School H has experienced continuous proficiency gains for all students in reading/language arts (+13.40%) and an overall gain in mathematics (+18.05%) between SY2009 and SY2011. There was an increase of 5.48% in reading/ language arts and essentially no change in mathematics (-.10%) during the initial SIG implementation year. As a point of contrast, School U has experienced an overall increase of 19.65% in reading/language arts and has an increase of 9.69% in mathematics between SY2009 and SY2011. The school has increased in reading/language arts proficiency (+11.58%), and, like School H, has remained relatively consistent in mathematics (-.43%) between SY2010 and SY2011. See Table 31 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 39. Achievement Profile for School H: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

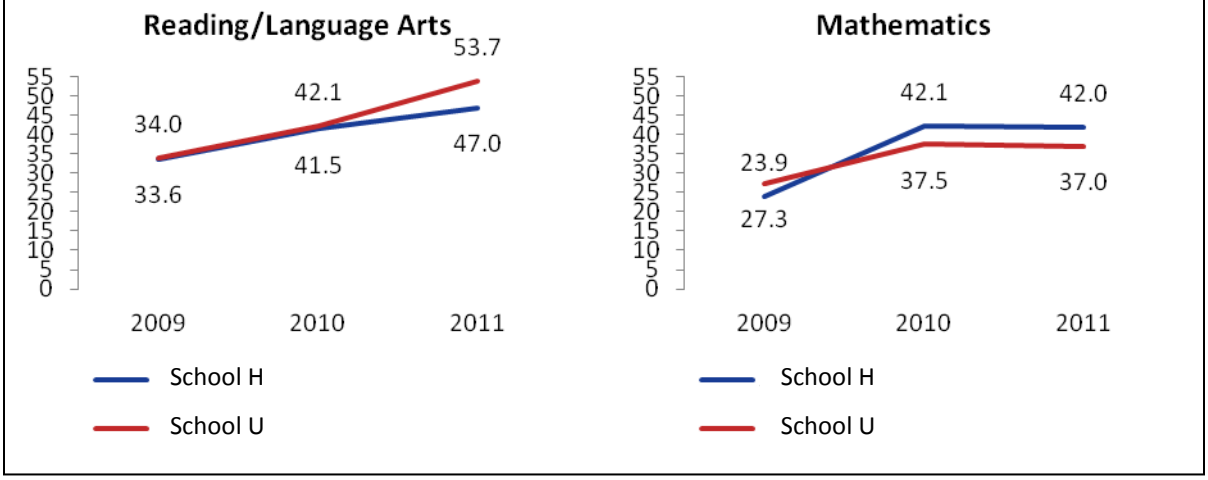


Table 31. Reading/language arts and Mathematics Gains for All Students in School H and School U

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School H	7.92	5.48	18.15	-0.10
School U	8.07	11.58	10.12	-0.43
School Δ	-0.15	-6.10	8.03	0.33

School O

Figure 40 indicates that School O has experienced continuous proficiency gains for all students in reading/language arts (+8.68%) and in mathematics (+3.59%) between SY2009 and SY2011. There was an increase of 8.31% in reading/language arts and an increase of 2.28% in mathematics during the initial SIG implementation year. As a point of contrast, School V has experienced essentially no net change in reading/language arts (+.85%) and a decrease of -2.58% in mathematics between SY2009 and SY2011. The school has experienced very little change in reading/language arts proficiency (-.59%) and mathematics (-.88%) between SY2010 and SY2011. See Table 32 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 40. Achievement Profile for School O: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

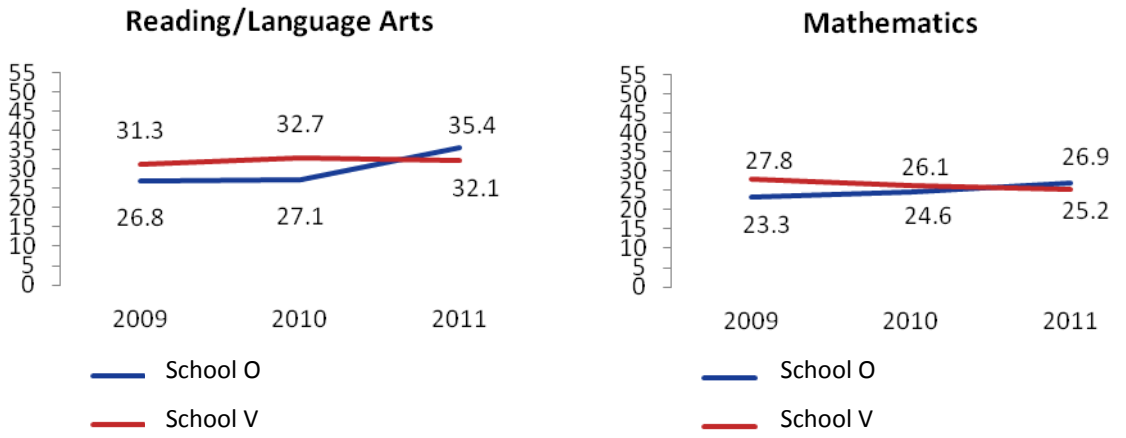


Table 32. Reading/Language Arts and Mathematics Gains for All Students in School O and School V

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School O	0.37	8.31	1.32	2.28
School V	1.44	-0.59	-1.70	-0.88
School Δ	-1.08	8.91	3.02	3.16

School L

Figure 41 indicates that School L has experienced net proficiency gains for all students in reading/language arts (+7.60%) and in mathematics (+3.47%) between SY2009 and SY2011. There was an increase of 8.71% in reading/language arts and a decrease of 2.70% in mathematics during the initial SIG implementation year. As a point of contrast, School W has experienced continuous improvement in reading/language arts (+15.27%) and in mathematics (+12.00%) between SY2009 and SY2011. The school has experienced a higher rate of change in reading/language arts proficiency (+10.44%) and mathematics (+1.89%) than School L between SY2010 and SY2011. See Table 33 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 41. Achievement Profile for School L: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

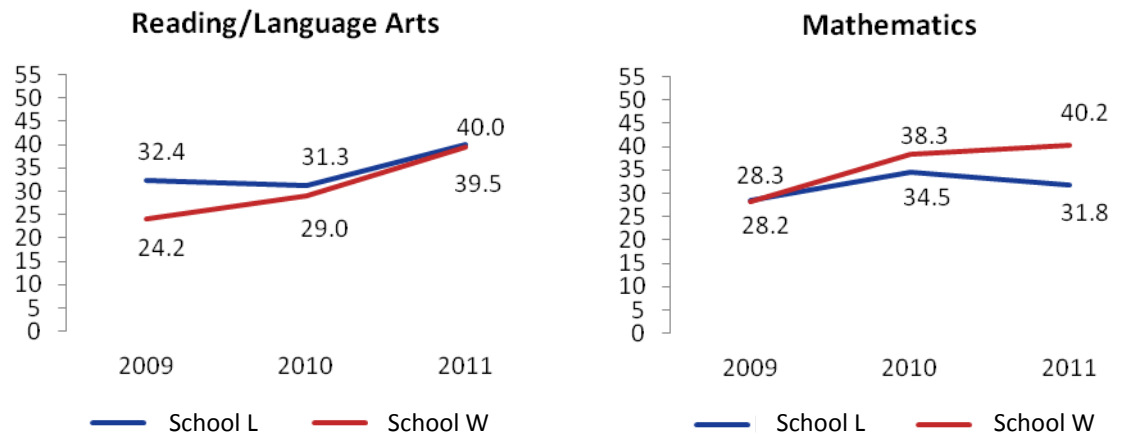


Table 33. Reading/Language Arts and Mathematics Gains for All Students in School L and School W

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School L	-1.11	8.71	6.17	-2.70
School W	4.83	10.44	10.12	1.89
School Δ	-5.94	-1.73	-3.95	-4.58

School J

Figure 42 indicates that School J has experienced continuous and rather noteworthy proficiency gains for all students in reading/language arts (+26.16%) and in mathematics (+24.23%) between SY2009 and SY2011. There was an increase of 13.25% in reading/language arts and an increase of 9.45% in mathematics during the initial SIG implementation year. The contrast school, School X was not tested in 2009 due to special circumstances. However, between SY2010 and SY2011, the school has experienced a decrease in reading/language arts (-2.57%) and a decrease in mathematics (-11.66). See Table 34 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 42. Achievement Profile for School J: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

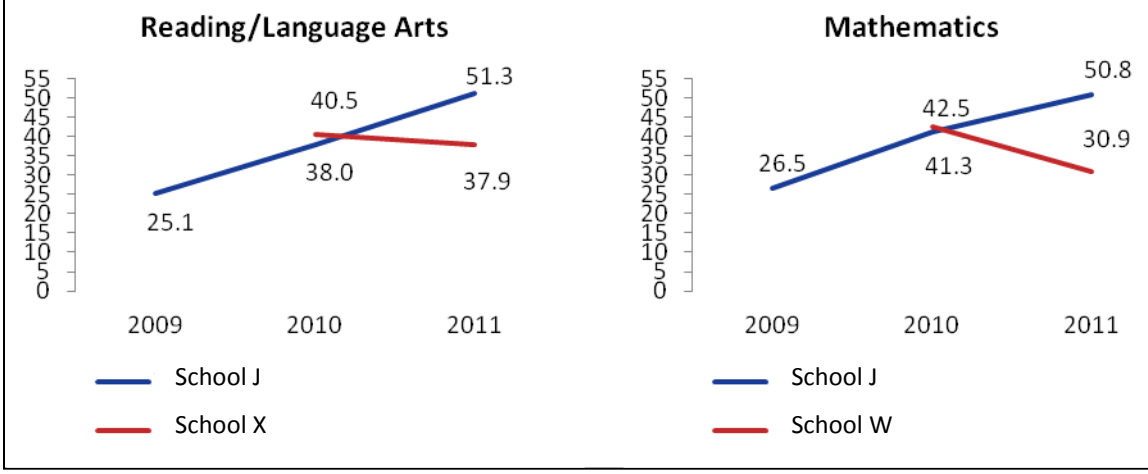


Table 34. Reading/Language Arts and Mathematics Gains for All Students in School J and School X

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School J	12.91	13.25	14.77	9.45
School X	NOT TESTED SY2009	-2.57	NOT TESTED SY2009	-11.66
School Δ		15.82		21.11

School C

Figure 43 indicates that School C has experienced continuous and noteworthy proficiency gains for all students in reading/language arts (+22.51%) and mathematics (+18.52%) between SY2009 and SY2011. There was an increase of 13.98% in reading/language arts and an increase of 9.14% in mathematics during the initial SIG implementation year. As a point of contrast, School Y has experienced a smaller net change in reading/language arts (+5.41%) and a continuous decrease of -9.35% in mathematics between SY2009 and SY2011. The school has experienced positive change in reading/language arts (+9.46%) and a decline of -6.02% in mathematics between SY2010 and SY2011. See Table 35 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 43. Achievement Profile for School C: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

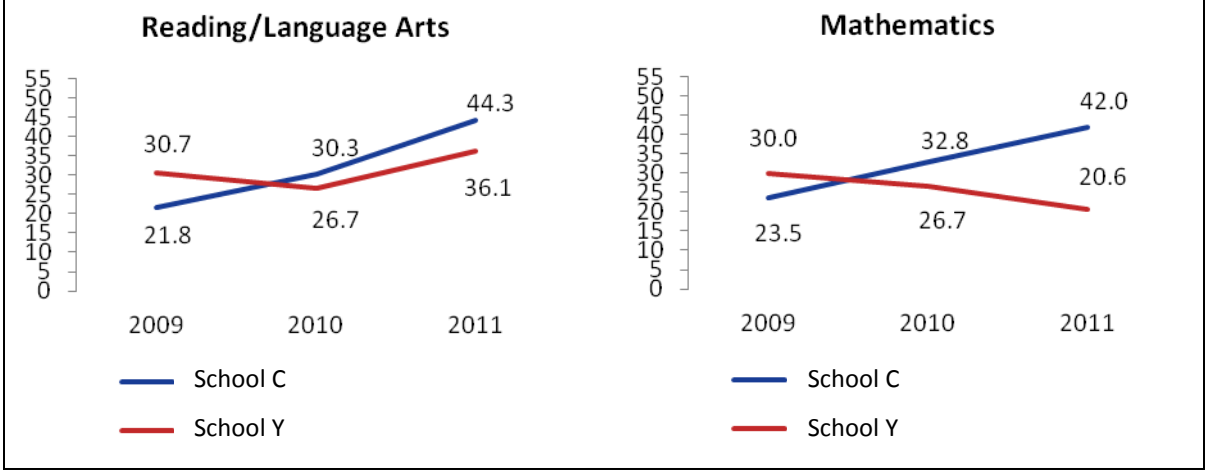


Table 35. Reading/Language Arts and Mathematics Gains for All Students in School C and School Y

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School C	8.52	13.98	9.38	9.14
School Y	-4.05	9.46	-3.33	-6.02
School Δ	12.57	4.52	12.71	15.16

School A

Figure 44 indicates that School A has experienced continuous proficiency gains for all students in reading/language arts (+12.89%) and a net increase in mathematics (+14.47%) between SY2009 and SY2011. There was an increase of 11.71% in reading/language arts and an increase of 14.98% in mathematics during the initial SIG implementation year. As a point of contrast, School Z has experienced large and continuous net gains in reading/language arts (+26.99%) and a sizable net gain of +15.98% in mathematics between SY2009 and SY2011. The school has experienced a large increase in reading/language arts proficiency (+16.19%) and a smaller gain in mathematics (+5.17%) between SY2010 and SY2011. See Table 36 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 44. Achievement Profile for School A: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

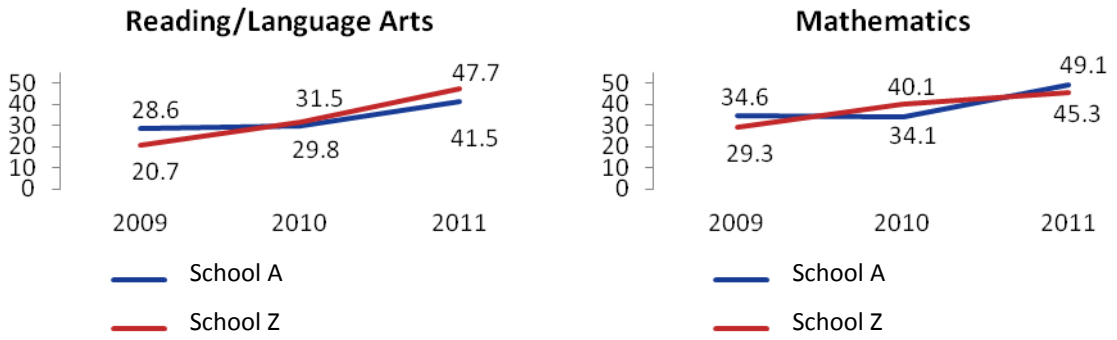


Table 36. Reading/Language Arts and Mathematics Gains for All Students in School A and School Z

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School A	1.18	11.71	-0.51	14.98
School Z	10.80	16.19	10.81	5.17
School Δ	-9.62	-4.48	-11.32	9.81

School M

Figure 45 indicates that School M has experienced continuous incremental proficiency gains for all students in mathematics (+5.20%) and a net gain in reading/language arts (+8.23%) between SY2009 and SY2011. There was very little change in reading/language arts (+.47%) and an increase of 1.39% in mathematics during the initial SIG implementation year. As a point of contrast, School AA has experienced a net change in reading/language arts (+5.09%) and a similar increase of +5.95% in mathematics between SY2009 and SY2011. The school has experienced a large increase in reading/language arts proficiency (+11.23%) and smaller gain in mathematics (+3.76%) between SY2010 and SY2011. See Table 37 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 45. Achievement Profile for School M: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

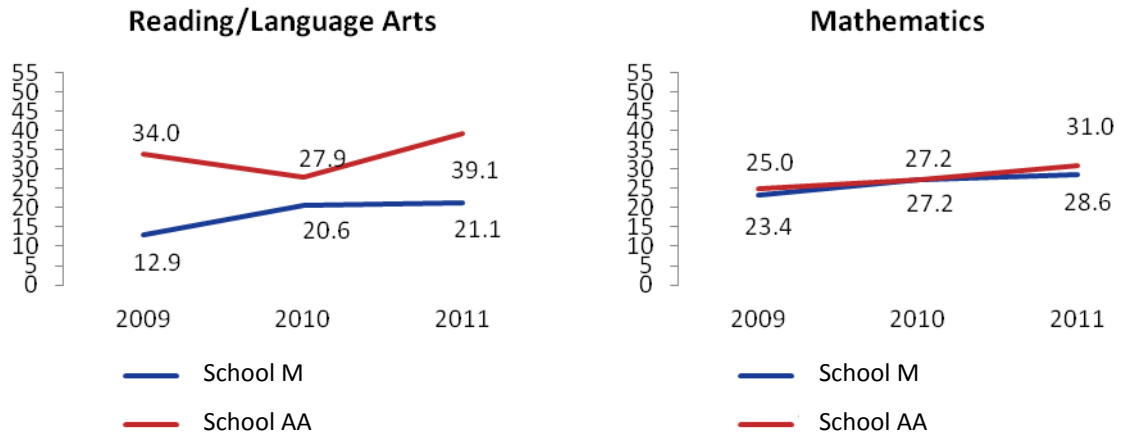


Table 37. Reading/Language Arts and Mathematics Gains for All Students in School M and School AA

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School M	7.76	0.47	3.82	1.39
School AA	-6.14	11.23	2.19	3.76
School Δ	13.90	-10.76	1.63	-2.38

School D

Figure 46 indicates that School D has experienced continuous proficiency gains for all students in reading/language arts (+17.86%) and in mathematics (+7.65%) between SY2009 and SY2011. There was an increase of 14.38% in reading/language arts and an increase of 3.60% in mathematics during the initial SIG implementation year. As a point of contrast, School AB has experienced a smaller net increase in reading/language arts (+6.45%) and similar increase of +7.07% in mathematics between SY2009 and SY2011. The school has experienced a small positive change in reading/language arts proficiency (+1.74%) and mathematics (+4.16%) between SY2010 and SY2011. See Table 38 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 46. Achievement Profile for School D: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

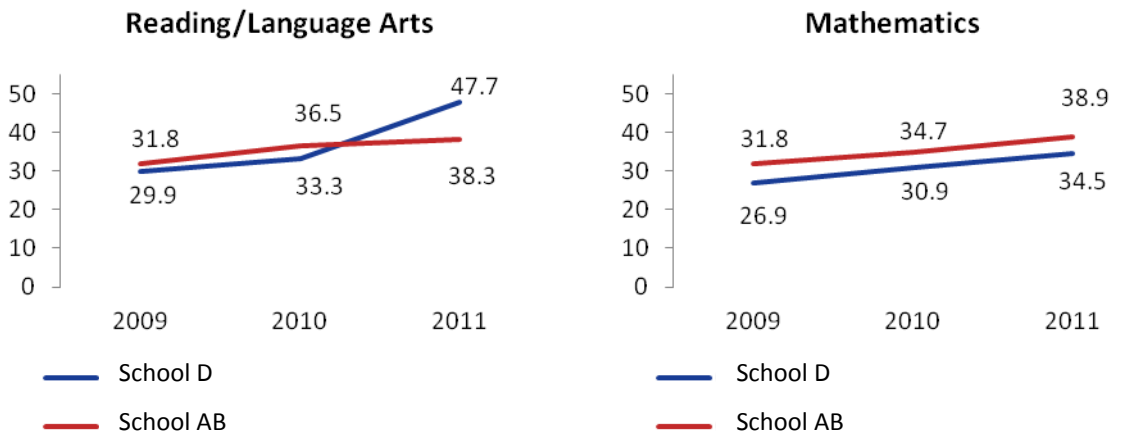


Table 38. Reading/Language Arts and Mathematics Gains for All Students in School D and School AB

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School D	3.48	14.38	4.05	3.60
School AB	4.71	1.74	2.91	4.16
School Δ	-1.23	12.64	1.14	-0.56

School F

Figure 47 indicates that School F has experienced continuous proficiency gains for all students in mathematics (+13.88%) and a net gain in reading/language arts (+8.40%) between SY2009 and SY2011. There was no substantial change in reading/language arts (-.91%) and an increase of 5.22% in mathematics during the initial SIG implementation year. As a point of contrast, School AC has experienced essentially a large net change in reading/language arts (+16.75%) and a considerable increase of +8.58% in mathematics between SY2009 and SY2011. The school has experienced a major increase in reading/language arts proficiency (+17.58%) and very little change in mathematics (+.64%) between SY2010 and SY2011. See Table 39 for a comparison of gains by school year. For an examination of subgroup proficiency rates, please see Appendix D.

Figure 47. Achievement Profile for School F: Comparison of Percent Proficient (SIG versus Non-SIG Comparison School), 2009–2011

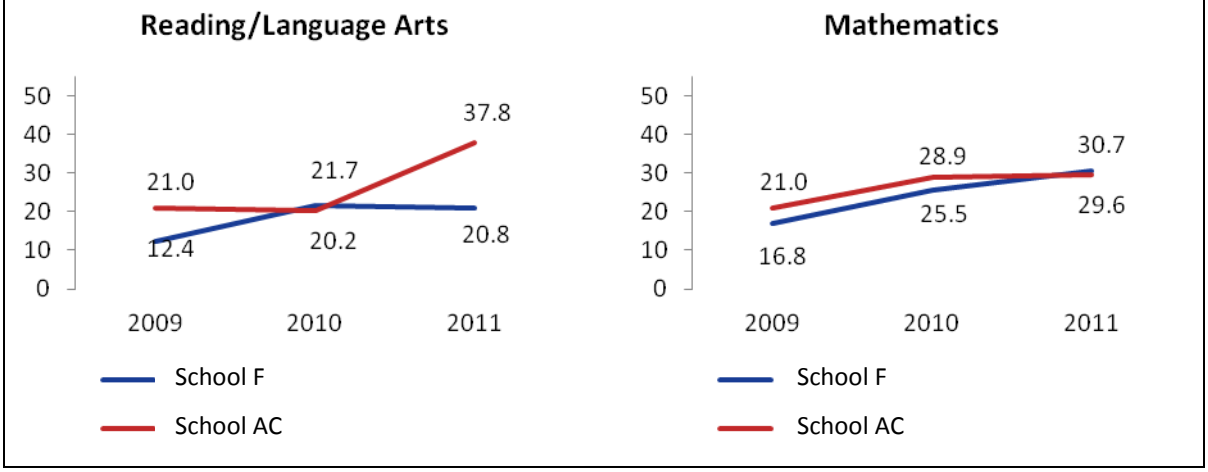


Table 39. Reading/Language Arts and Mathematics Gains for All Students in School F and School AC

	Reading/Language Arts Gains		Mathematics Gains	
School	SY2009 – SY2010	SY2010 – SY2011	SY2009 – SY2010	SY2010 – SY2011
School F	9.31	-0.91	8.66	5.22
School AC	-0.83	17.58	7.94	0.64
School Δ	10.14	-18.49	0.72	4.58

Evaluation Question 4

EQ4. To what extent have SIG schools successfully implemented and institutionalized practices and structures that are supportive of continuous school improvement?

To investigate this question, we analyzed data collected by Title I staff, using two additional surveys of educators working in SIG schools—the Measure of School Capacity (MSCI) for Improvement and the School Culture Survey.

Measure of School Capacity for Improvement results

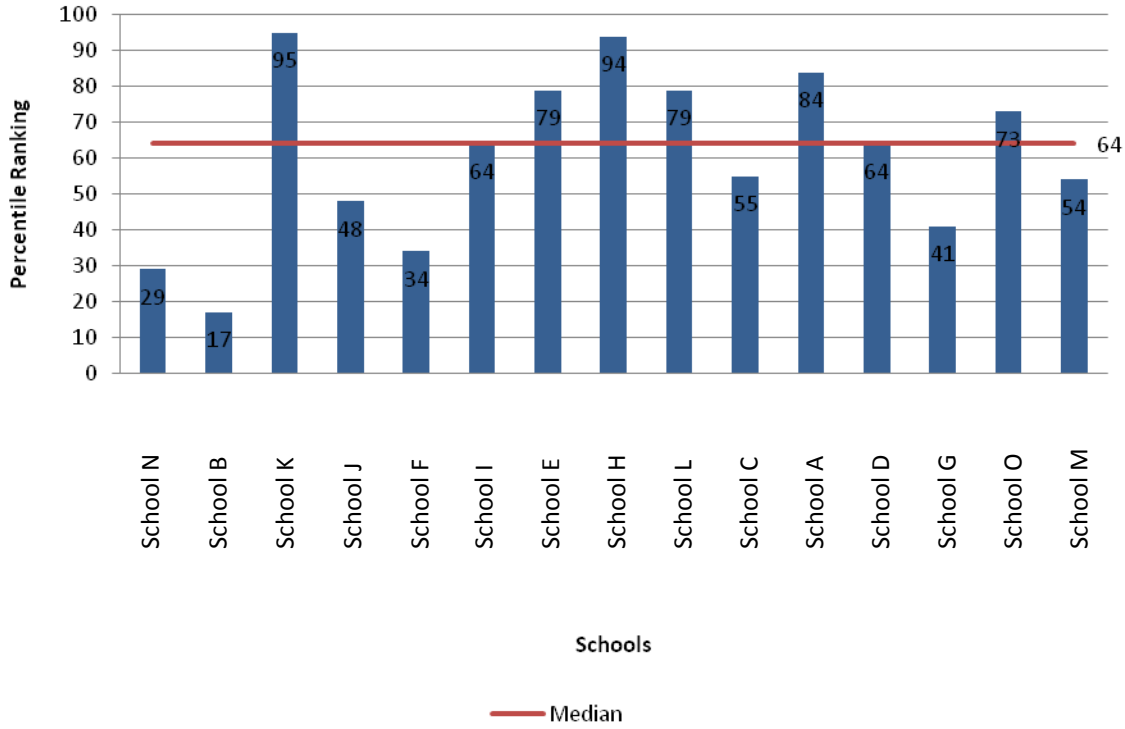
This section presents the results of the MSCI survey for the school year 2010-2011.

Figure 48 through Figure 54 illustrate the percentile ranking for each of the 15 SIG schools. These percentile ranks are derived by comparing the mean subscale score for each SIG school with data from a larger group of norm-referenced schools in the *MSCI User Manual and Technical Report* (Hughes, et al., 2006). For this analysis, we compared each SIG school to schools of the same programmatic level (e.g., elementary, middle, and high). Each bar represents the percentile rank for an individual SIG school. The line that crosses the bars on each chart represents the median percentile rank for all SIG schools. The median percentile serves as a summary measure of the rank of the SIG schools when they are aggregated into a single group. If the median for SIG schools falls below the 50th percentile, this would mean that the SIG schools exhibited an MSCI subscale score that was lower than 50% of the norm-referenced schools.

Equity in practice

The median percentile rank of SIG schools on this subscale was 64 (Figure 48), which means 36% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *fourth-highest ranking* of the MSCI subscale scores, indicating it as an area of relative strength for SIG schools. With respect to individual results, School B exhibited the lowest percentile ranking on the subscale, scoring lower than 83% of norm-referenced schools. School K and School H exhibited the highest rankings, with only 5% and 6% of norm-referenced schools scoring higher, respectively.

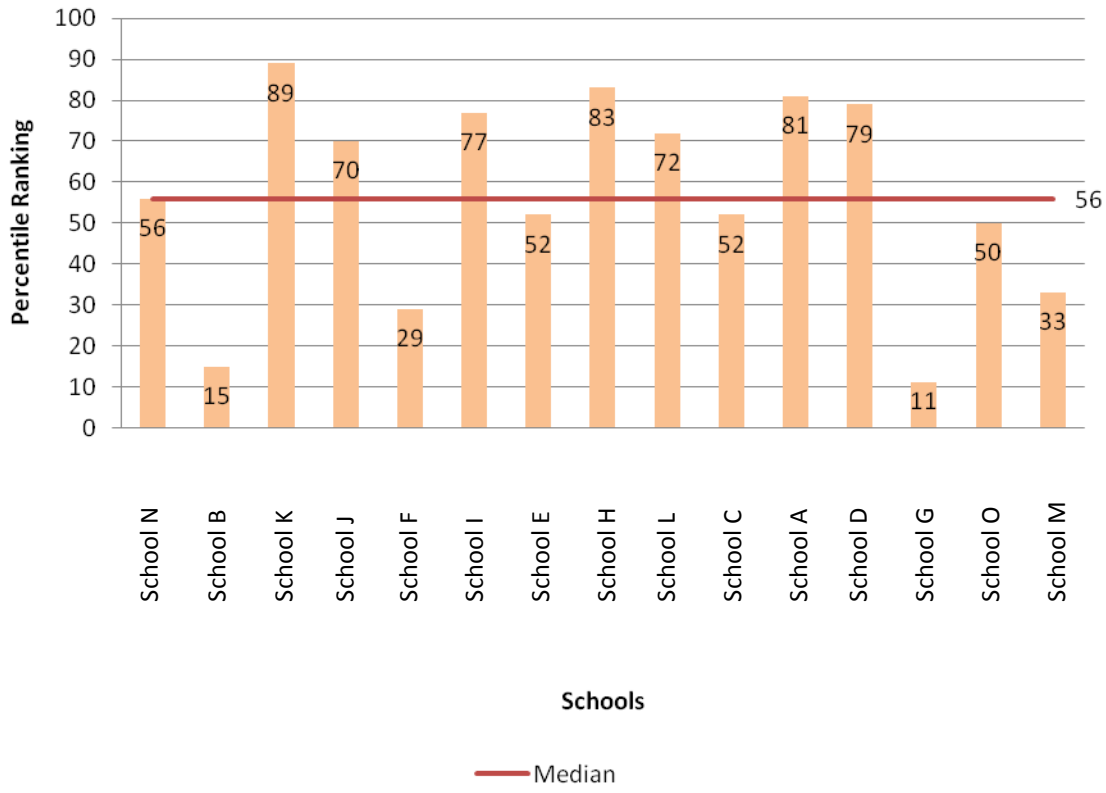
Figure 48. Equity in Practice: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Expectations for student performance

The median percentile rank of SIG schools on this subscale was 56 (Figure 49), which means 44% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *second-lowest ranking* of the MSCI subscales, indicating it as an area of concern for SIG schools. With respect to individual results, School G and School B exhibited the lowest percentile ranks on the subscale, scoring lower than 89% and 85% of norm-referenced schools, respectively. School K exhibited the highest ranking, with only 11% of norm-referenced schools scoring higher.

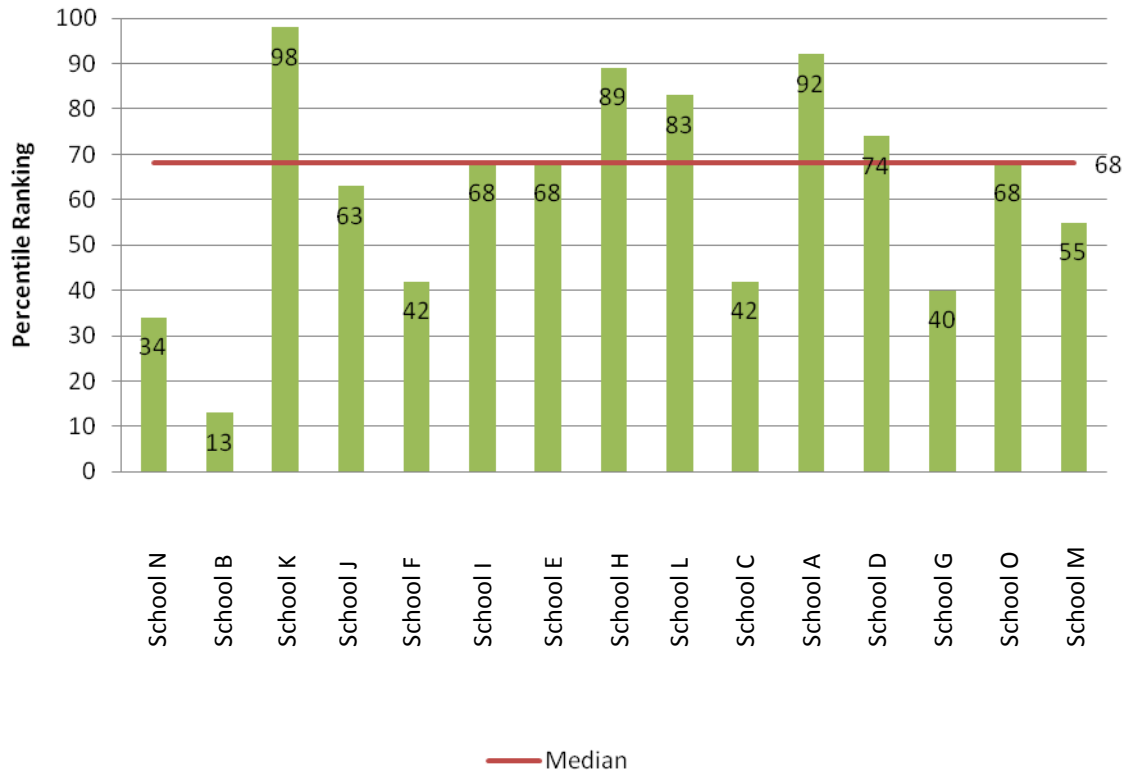
Figure 49. Expectations for Student Performance: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Differentiated instruction

The median percentile rank of SIG schools on this subscale was 68 (Figure 50), which means 32% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *third-highest ranking* of the MSCI subscales, indicating it as an area of relative strength for SIG schools. With respect to individual results, School B exhibited the lowest percentile rank on the subscale, scoring lower than 87% of norm-referenced schools. School K and School A exhibited the highest ranking, with only 2% and 8% of norm-referenced schools scoring higher respectively.

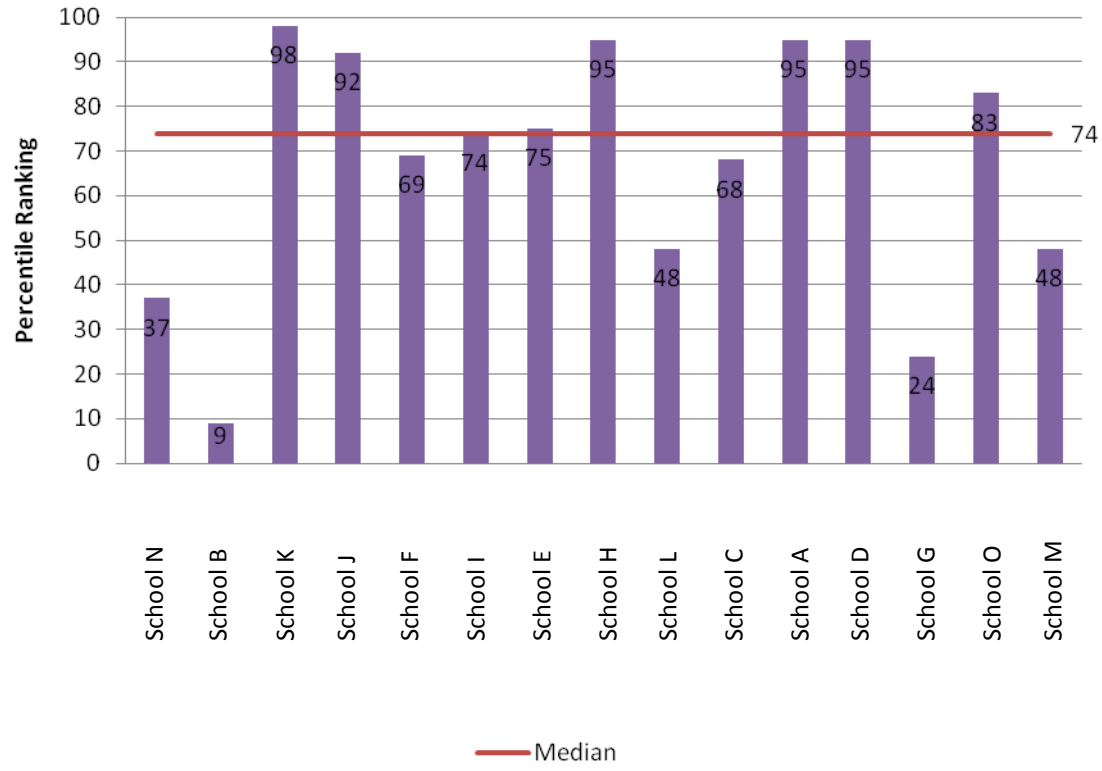
Figure 50. Differentiated Instruction: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Improvement program coherence

The median percentile rank of SIG schools on this subscale was 74 (Figure 51), which means 26% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *highest ranking* of the MSCI subscales, indicating it as an area of particular strength for SIG schools. With respect to individual results, School B exhibited the lowest percentile ranking on the subscale, scoring lower than 91% of norm-referenced schools. School K exhibited the highest ranking, with only 2% of norm-referenced schools scoring higher. Three other schools, School H, School A, and School D were also ranked high, with only 5% of norm-referenced schools scoring higher than these schools.

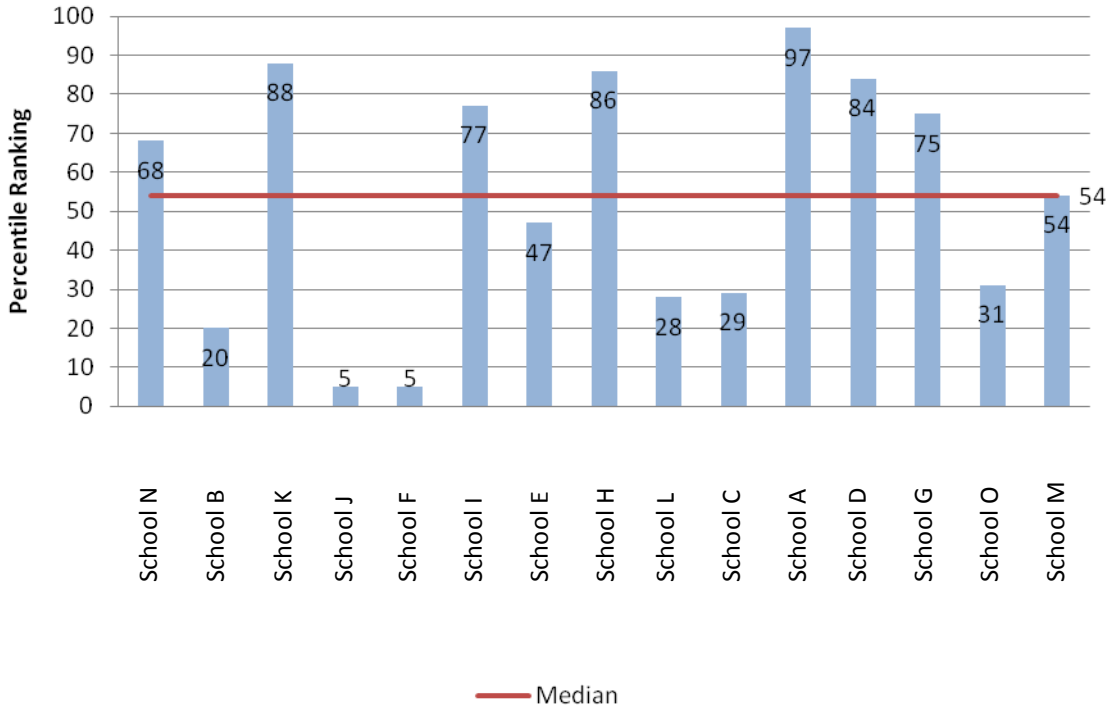
Figure 51. Improvement Program Coherence: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Peer reviewed practice

The median percentile rank of SIG schools on this subscale was 54 (Figure 52), which means 46% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *lowest ranked* of the MSCI subscales, indicating it as an area of particular concern for SIG schools. With respect to individual results, School J and School F exhibited the lowest percentile ranks on the subscale, scoring lower than 95% of norm-referenced schools. School A exhibited the highest rank, with only 3% of norm-referenced schools scoring higher.

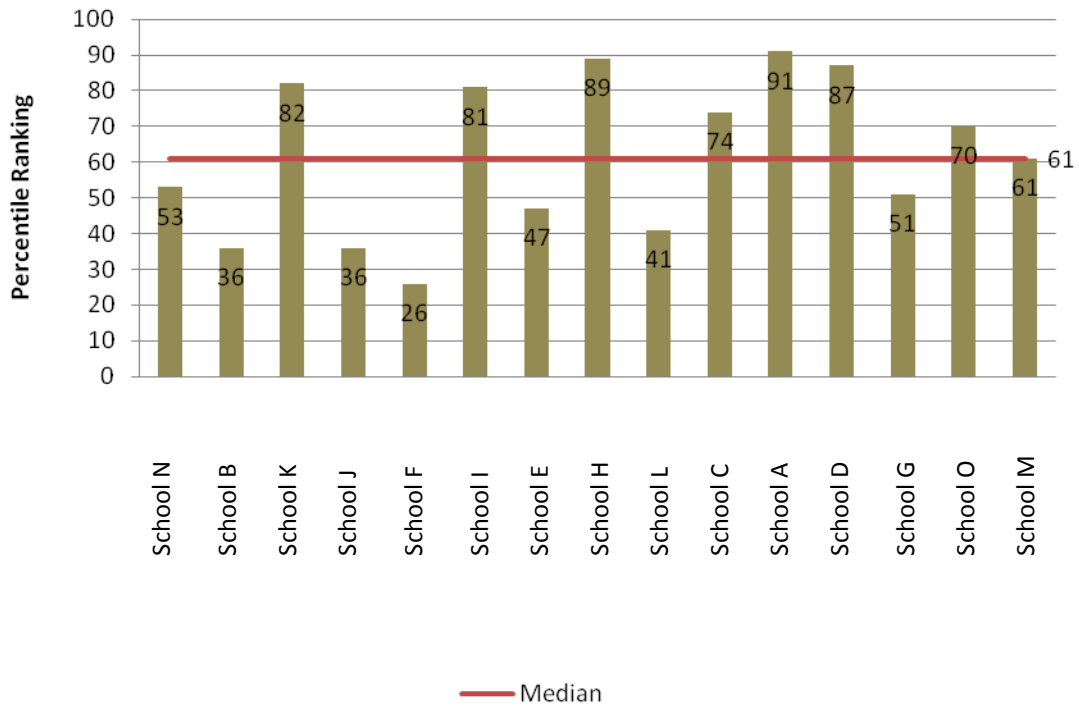
Figure 52. Peer Reviewed Practice: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Coordinated curriculum

The median percentile rank of SIG schools on this subscale was 61 (Figure 53), which means that 39% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *third-lowest rated* of the MSCI subscales, indicating it as an area of concern for SIG schools. With respect to individual results, School F exhibited the lowest percentile ranks on the subscale, scoring lower than 74% of norm-referenced schools. School A exhibited the highest ranking, with only 9% of norm-referenced schools scoring higher.

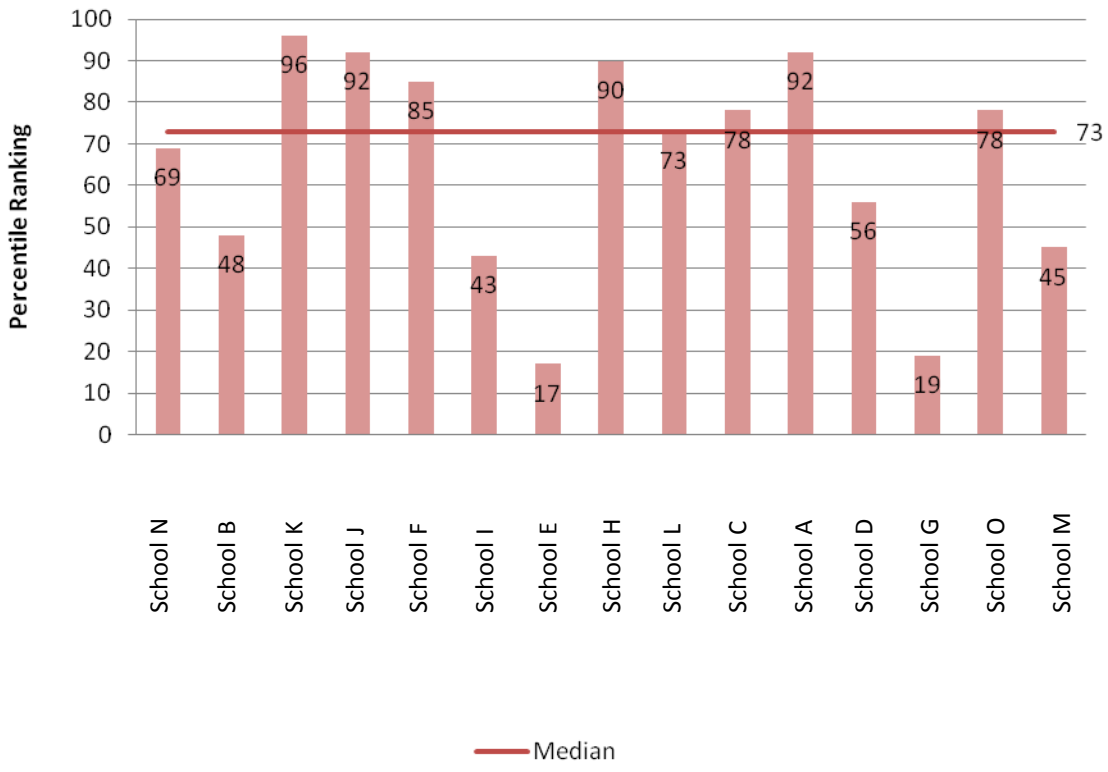
Figure 53. Coordinated Curriculum: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



Technical resources

The median percentile rank of SIG schools on this subscale was 73 (Figure 54), which means that 27% of norm-referenced schools exhibited scores on the subscale that were higher than the scores of the SIG schools. This was the *second-highest ranking* of the MSCI subscales, indicating it as an area of particular strength for SIG schools. Four schools exhibited percentile ranks at or above 90 (School K [96], School J [92], School A [92], and School H [90]). With respect to individual results, School E and School G exhibited the lowest percentile ranking on the subscale, scoring lower than 83% and 81% of norm-referenced schools respectively.

Figure 54. Technical Resources: Percentile Rank of Individual SIG Schools on MSCI Norm-Referenced Scale, and Median Score for All SIG Schools, 2010-2011



School Culture Survey results

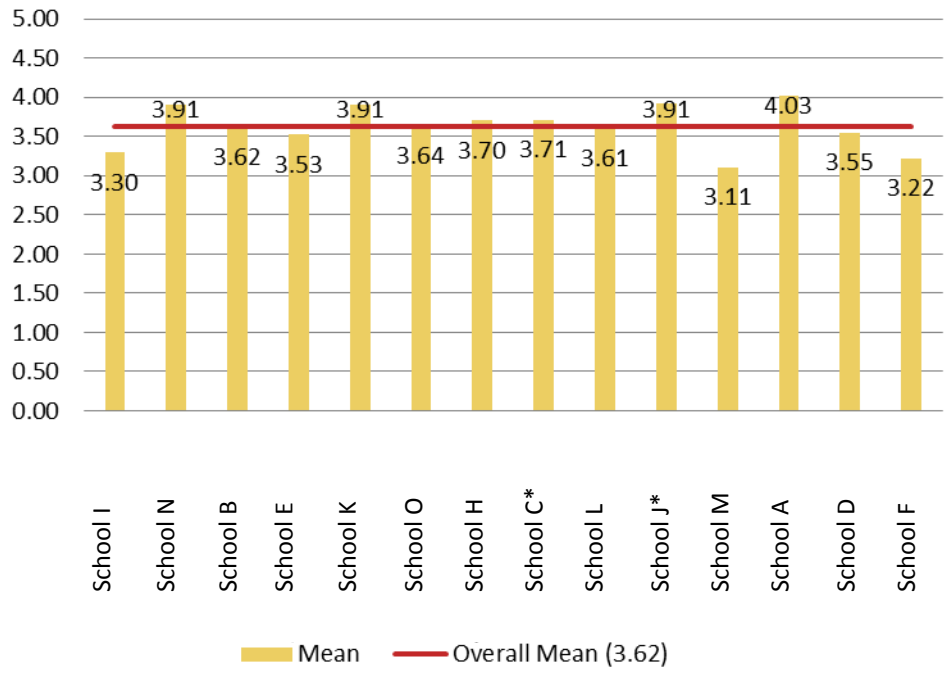
The School Culture Survey uses a 5-point Likert-type response scale (i.e., 1 = S*trongly disagree*, 2 = *Disagree*, 3 = *Neutral*, 4 = *Agree*, 5 = *Strongly agree*). The midpoint of the scale, a neutral response (3.00), is used in this report as a point of reference for interpreting data from the survey. Any mean score below 3.00 would indicate general disagreement with survey items and any mean score above 3.00 would indicate general agreement with the item. Thus, a higher mean score (i.e., 4.00 or above) indicates that professional staff has a favorable perception regarding their school culture.

Overall average score by domain

On a scale of 1 to 5, the overall average rating for all schools on all seven domains was 3.62, with little variation among the domains (SD = 0.34) (Table 40). Mean domain scores for all schools ranged between 3.04 (Learning Partnership) and 3.95 (Professional Development). In addition to the Learning Partnership domain, two additional domains, Collaborative Leadership (3.60) and Teacher Collaboration (3.35), had mean scores below the overall average. The remaining four domains had average scores that were above the overall mean.

Figure 55 depicts overall mean scores for each school while Table 41 provides overall mean scores for each school. The mean score for the seven domains among schools ranged between 3.11 for School M and 4.03 for School A (SD = 0.36), with 42.9% of the schools (n = 6) scoring below the group average (M = 3.62) and none of the schools scoring below 3.00, the midpoint of the scale (Figure 1; Table 3). The greatest amount of variation was found in the Collaborative Leadership domain, with a lowest score of 2.77 and highest score of 4.14 (SD = 0.45). For the lowest scoring domain, *Learning Partnership*, 64.3% of the schools (n = 9) had average scores below the midpoint of the scale (Table 4). Findings *clearly indicate that this may be an area in need of more focused attention.*

Figure 55. Overall Mean School Culture Survey Scores by School



*Data from SY2009-2010. Data not available for SY2010-2011.

Table 40. Overall Mean School Culture Survey Scores by Domain

	Collaborative Leadership	Teacher Collaboration	Professional Development	Unity of Purpose	Collegial Support	Learning Partnership	Efficacy Factor	Overall Mean* (SD)
Mean	3.60	3.35	3.95	3.90	3.91	3.04	3.63	3.62 (0.34)

*1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Table 41. Mean School Culture Survey Scores by Domain by School

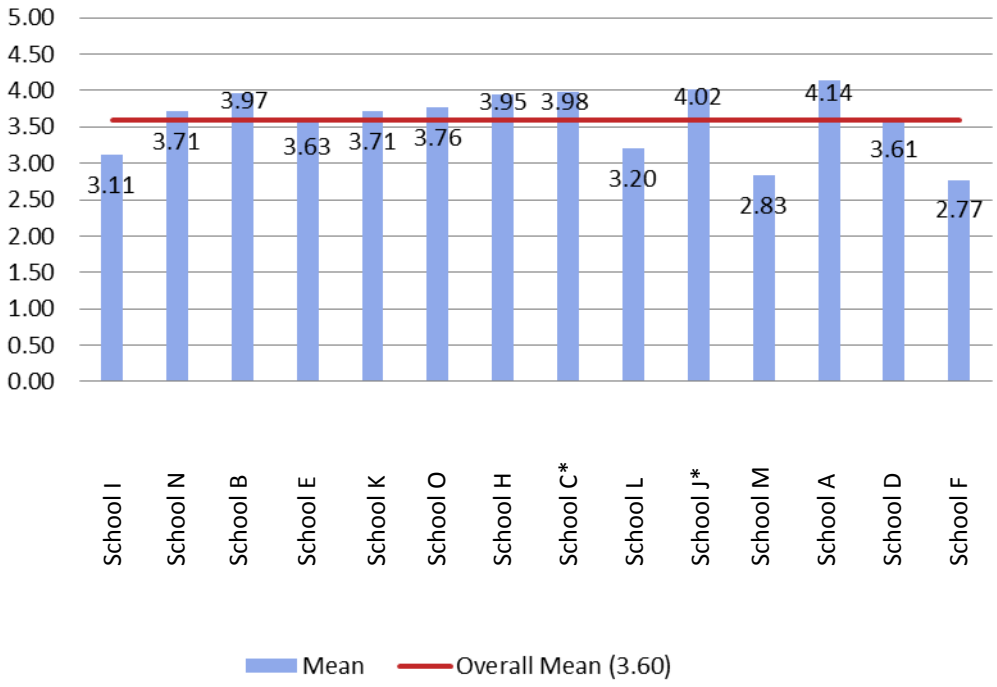
	Culture Survey Sub Scales								
School	Collaborative Leadership	Teacher Collaboration	Professional Development	Unity of Purpose	Collegial Support	Learning Partner-ship	Efficacy Factor	Mean**	SD
Mean (SD)	3.60 (0.45)	3.35 (0.43)	3.95 (0.28)	3.90 (0.29)	3.91 (0.27)	3.04 (0.37)	3.63 (0.21)	3.62	(0.34)
School I	3.11	**2.90**	3.56	3.59	3.55	**2.86**	3.50	3.30	0.33
School N	3.71	3.66	4.25	4.12	4.12	3.58	3.95	3.91	0.26
School B	3.97	3.45	3.85	3.88	3.86	**2.89**	3.43	3.62	0.39
School E	3.63	3.32	3.89	3.85	3.88	**2.86**	3.26	3.53	0.39
School K	3.71	3.66	4.25	4.12	4.12	3.58	3.95	3.91	0.26
School O	3.76	3.26	4.11	3.87	4.07	**2.73**	3.65	3.64	0.49
School H	3.95	3.38	4.01	4.02	4.02	**2.92**	3.63	3.70	0.42
School C*	3.98	3.55	4.13	4.00	3.99	**2.81**	3.49	3.71	0.46
School L	3.20	3.13	3.91	3.96	3.93	3.29	3.84	3.61	0.38
School J*	4.02	3.79	4.12	4.21	3.95	3.62	3.69	3.91	0.22
School M	**2.83**	**2.71**	3.56	3.47	3.43	**2.38**	3.36	3.11	0.46
School A	4.14	4.27	4.39	4.33	4.50	**2.85**	3.70	4.03	0.58
School D	3.61	3.01	3.63	3.86	3.74	3.25	3.78	3.55	0.31
School F	**2.77**	**2.78**	3.64	3.29	3.61	**2.90**	3.58	3.22	0.40

*Data from SY2009-2010. Data not available for SY2010-2011.
**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree
Red indicates general disagreement with survey items (mean rating below 3.0).

Collaborative relationships

The average score for the 14 schools for the domain *Collaborative Relationships* was 3.60 (Figure 56, Table E- 1 in Appendix E, page 154) and was the *third lowest average score* among the seven domains. Scores ranged between 2.77 (School F) and 4.14 (School A) with a standard deviation of 0.45, which is the largest standard deviation amongst average item scores of the seven domains. All but two schools (n = 12) scored above the midpoint of the scale with 71.4% of the schools (n = 10) scoring above the group mean. The average score for each of the 11 items/indicators for the domain ranged between 3.26 and 4.03 (Table E- 1, page 154). The survey item measuring the extent to which teachers are rewarded for experimenting with new ideas and techniques received the lowest score while the item measuring the extent to which teachers are encouraged to share ideas received the highest score. Two other items that received lower scores should also be noted; the extent to which teachers are involved in the decision-making process (M = 3.37) and the degree to which teachers' involvement in policy or decision-making process is taken seriously (M = 3.39). Combined, scores for these three survey items suggest that *school administrations should make a better effort to empower teachers by encouraging their involvement in decision-making process and acknowledge their contributions*.

Figure 56. Mean School Culture Survey Scores for Collaborative Leadership, by School

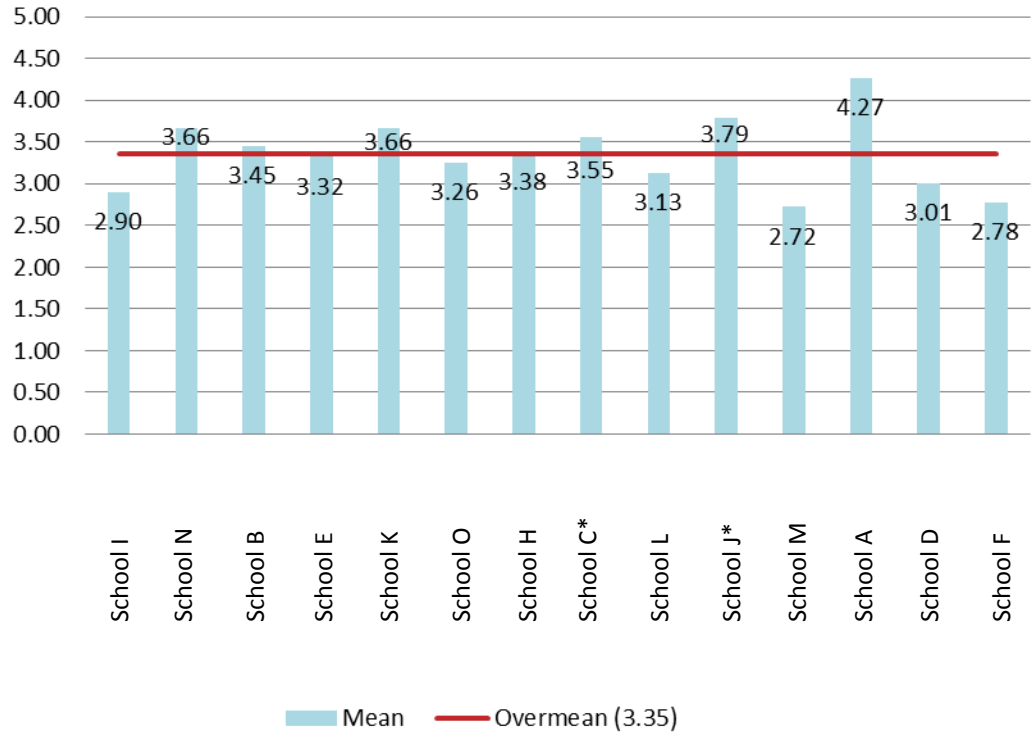


*Data from SY2009-2010. Data not available for SY2010-2011.

Teacher collaboration

For the domain *Teacher Collaboration*, the average domain score for all schools was 3.35, which was the *second lowest average score* among the seven domains (Figure 57, Table E- 2 in Appendix E, page 155). Average scores ranged between 2.72 (School M) and 4.27 (School A) with a standard deviation of 0.43, which was the second largest standard deviation among average scores for the seven domains. Seven out of fourteen schools (50%) had scores that were above the mean. Out of the seven schools with scores below the mean for this domain, three schools (School F, School I, and School M) scored below the midpoint of the scale (Figure 57), indicating that this may be an area of concern for these schools. Across all 14 schools, average scores for five out of the six items were above the midpoint of the scale (Table E- 2, page 155). The item measuring the extent to which teachers take time to observe each other teaching received the lowest score with 2.69, well below the midpoint of the scale. Only one school, School A (M = 4.10) had a score above the midpoint of the scale for this survey item. *This indicates that this may be an area in need of more focused attention for almost all SIG schools.*

Figure 57. Mean School Culture Survey Scores for Teacher Collaboration, by School

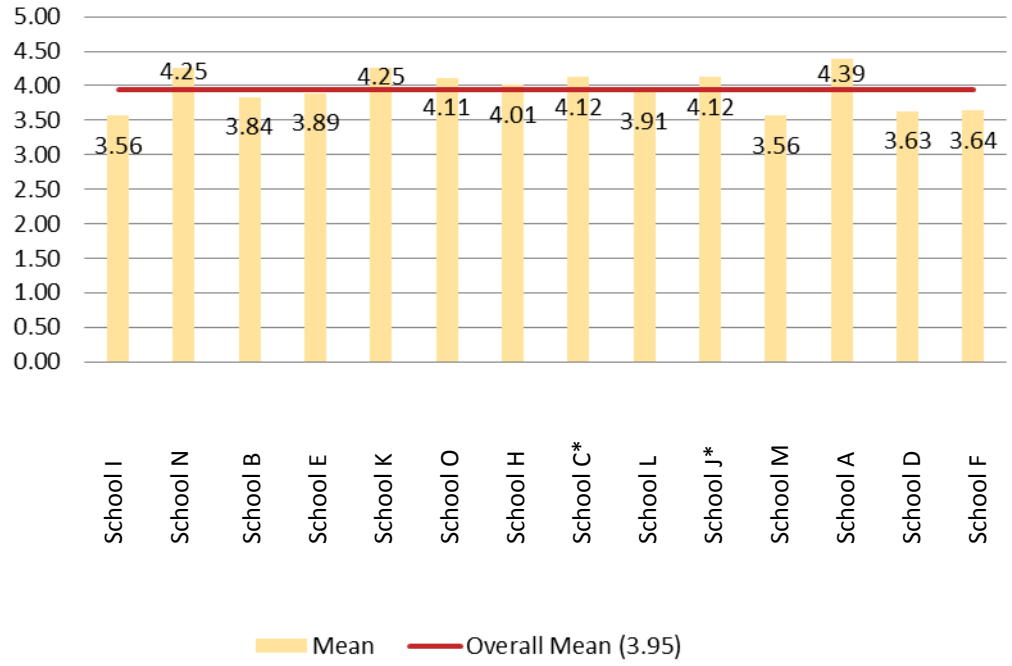


*Data from SY2009-2010. Data not available for SY2010-2011.

Professional development

The average score for the 14 schools for the domain *Professional Development* was 3.95 (Figure 58, Table E- 3 in Appendix E, page 156) and was the *highest average score* among the seven domains. Average scores ranged between 3.56 (School I and School M) and 4.39 (School A) with a standard deviation of 0.28. Seven out of fourteen schools (50%) had average scores above the mean for this domain (Figure 58). The average score for each of the 5 items/indicators for this domain ranged between 3.58 and 4.23 (Table E- 3, page 156). The item measuring the extent to which professional development was valued by the faculty received the lowest score while the item measuring the extent to which the faculty values school improvement received the highest score. *Overall, findings indicate this domain to be an area of particular strength for SIG schools.*

Figure 58. Mean School Culture Survey Scores for Professional Development, by School

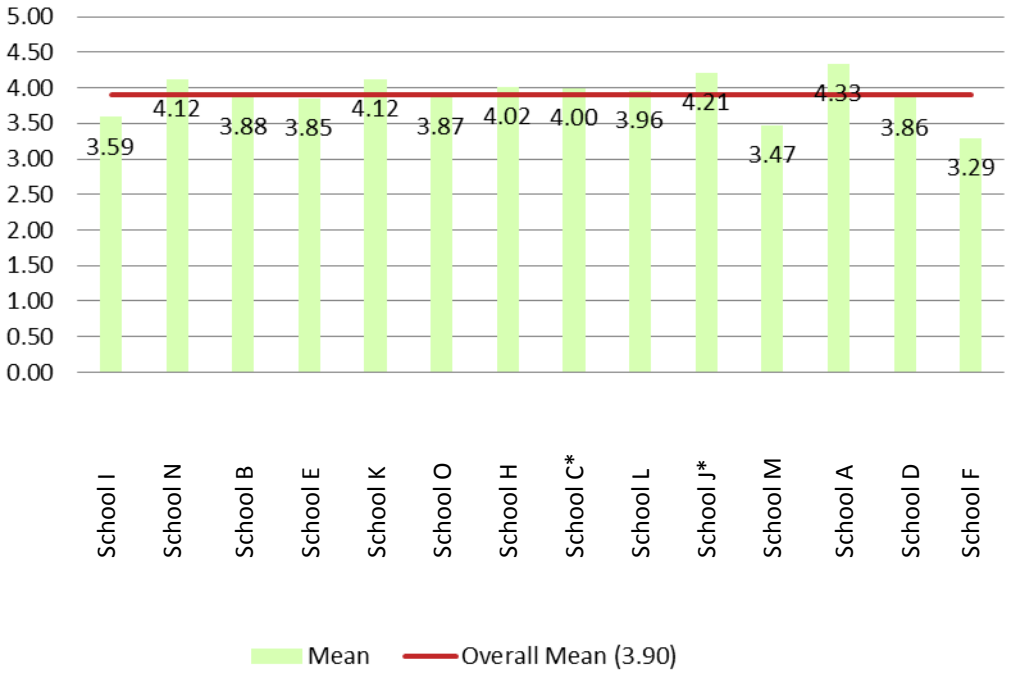


*Data from SY2009-2010. Data not available for SY2010-2011.

Unity of purpose

Mean scores for the domain *Unity of Purpose* ranged between 3.29 (School F) and 4.33 (School A) with an overall average score of 3.90 for all schools, which is the *third highest average score* among the seven dimensions (SD = 0.29) (Figure 59, Table E- 4 in Appendix E, page 157). Seven out of fourteen schools (50%) had mean scores below the overall average score for this domain. Average scores for the five indicators that made up this domain ranged between 3.53 for the extent to which the school mission statement reflects the values of the community and 4.07 for the extent with which teachers support the mission of the school (Table E- 4, page 157). *Findings indicate that this is an area of relative strength for SIG schools.*

Figure 59. Mean School Culture Survey Scores for Unity of Purpose, by School

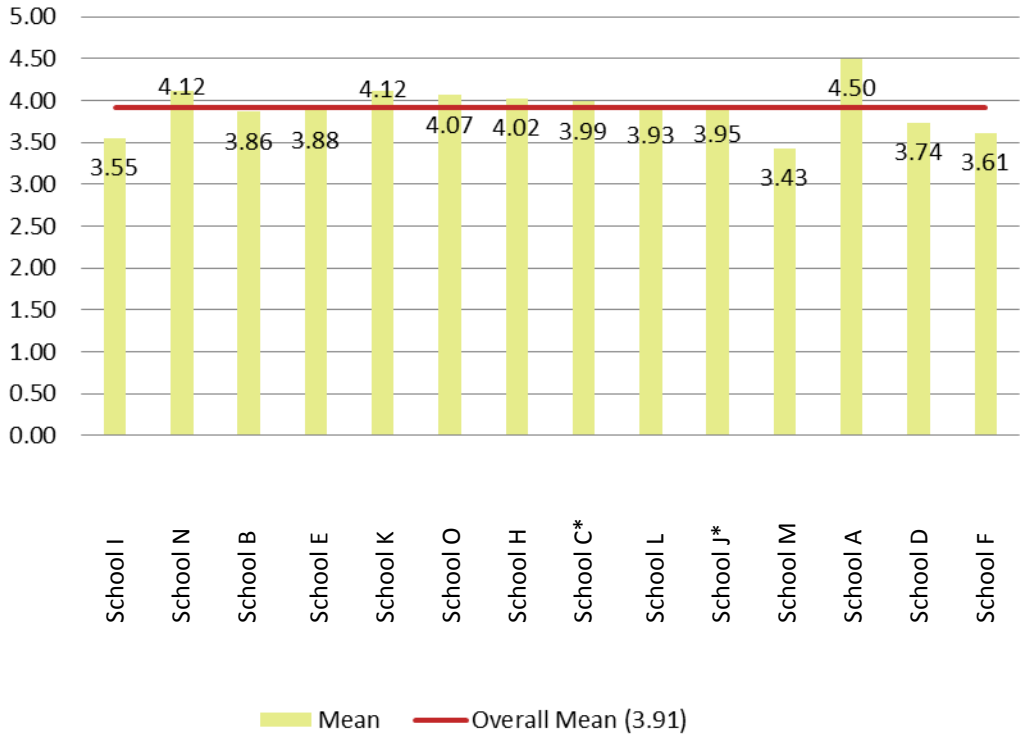


*Data from SY2009-2010. Data not available for SY2010-2011.

Collegial support

For the domain *Collegial Support*, the average domain score for all schools was 3.91, which was the *second highest average score* among the seven domains (Figure 60, Table E-5 in Appendix E, page 158). Average scores ranged between 3.43 (School M) and 4.50 (School A) with a standard deviation of 0.27. Eight out of fourteen schools (57.1%) had scores that were above the mean for the domain (Table E- 5, page 158). The item measuring the extent to which teachers trust each other received the lowest score (M = 3.60) and the degree to which teachers are willing to help out whenever there is a problem received the highest score (M = 4.19). *Findings indicate that this is an area of particular strength for SIG schools*.

Figure 60. Mean School Culture Survey Scores for Collegial Support, by School

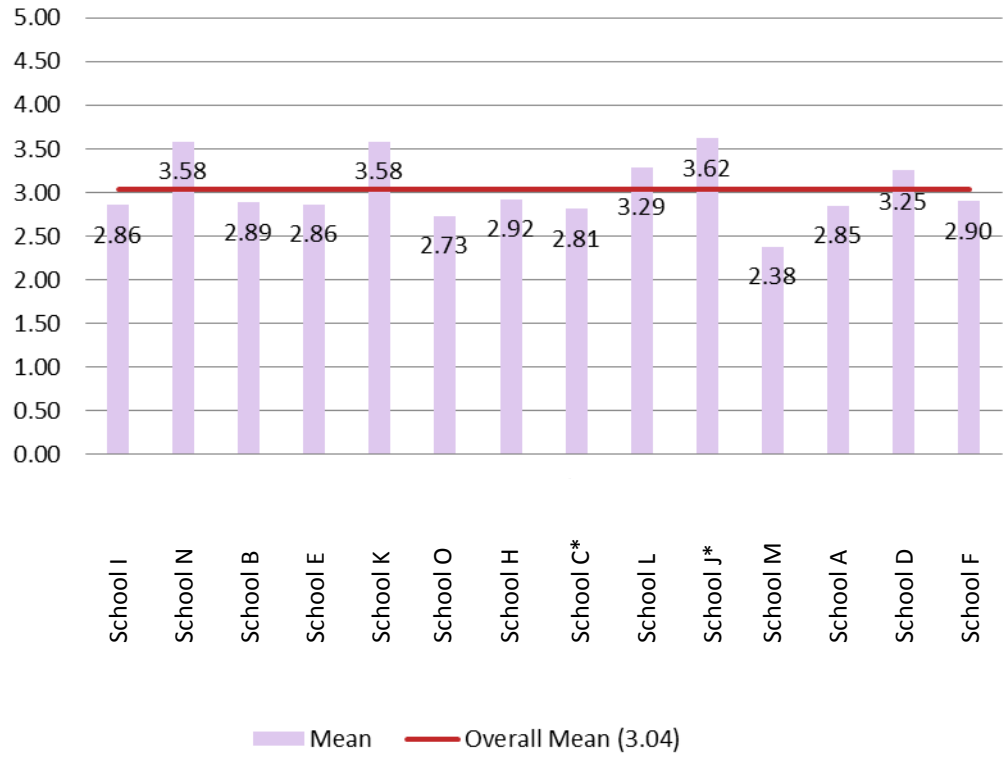


*Data from SY2009-2010. Data not available for SY2010-2011.

Learning partnerships

The average score for the 14 schools for the domain *Learning Partnership* was 3.04 (Figure 61, Table E- 6 in Appendix E, page 159) and was the *lowest average score* among the seven domains. Average scores ranged between 2.38 (School M) and 3.62 (School J) with a standard deviation of 0.37. Only 35.7% of schools (n = 5) scored above the average for the domain while 64.3 % of schools (n = 9) scored below the midpoint of the scale, *indicating this is an area of particular concern*. The average score for each of the 4 items/indicators for this domain ranged between 2.69 and 3.45 (Table E- 6, page 159). The item measuring the degree to which students accept responsibility for their schooling received the lowest score while the item measuring the frequency with which teachers and parents communicate about student performance received the highest score.

Figure 61. Mean School Culture Survey Scores for Learning Partnerships, by School

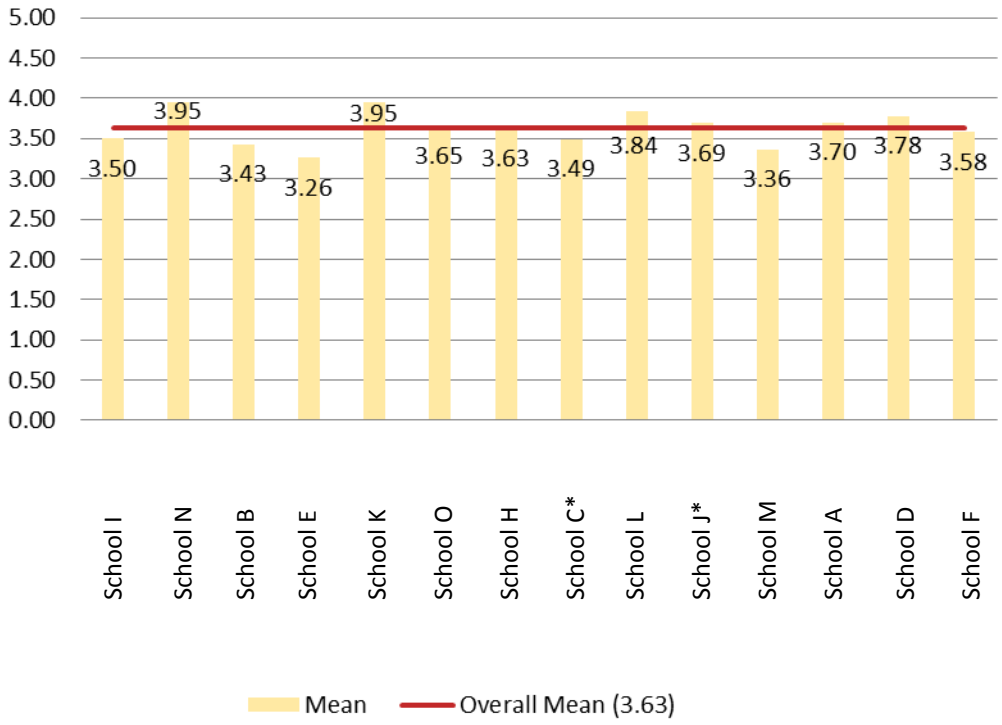


*Data from SY2009-2010. Data not available for SY2010-2011.

Efficacy factor

For the domain *Efficacy Factor*, the average domain score for all schools was 3.63, which was the *fourth lowest average score* among the seven domains (Figure 62, Table E- 7 in Appendix E, page 160). The domain average is only slightly better (one tenth of a point) than the overall mean for all seven domains. Average scores ranged between 3.26 (School E) and 3.95 (School N and School K) with a standard deviation of 0.21. Seven schools (50%) had scores which were above the group mean and all schools (n = 14) scored above the mid-point of the scale (Figure 62). The average score for each of the 10 items/indicators for this domain ranged between 2.91 and 4.27 (Table E- 7, page 160). The survey item measuring the extent to which faculty believes that the attitudes and habits students bring to class greatly reduce their chances for academic success received the lowest score while the degree to which teachers are expected to help maintain discipline in the entire school received the highest score.

Figure 62. Mean School Culture Survey Scores for Efficacy Factor, by School

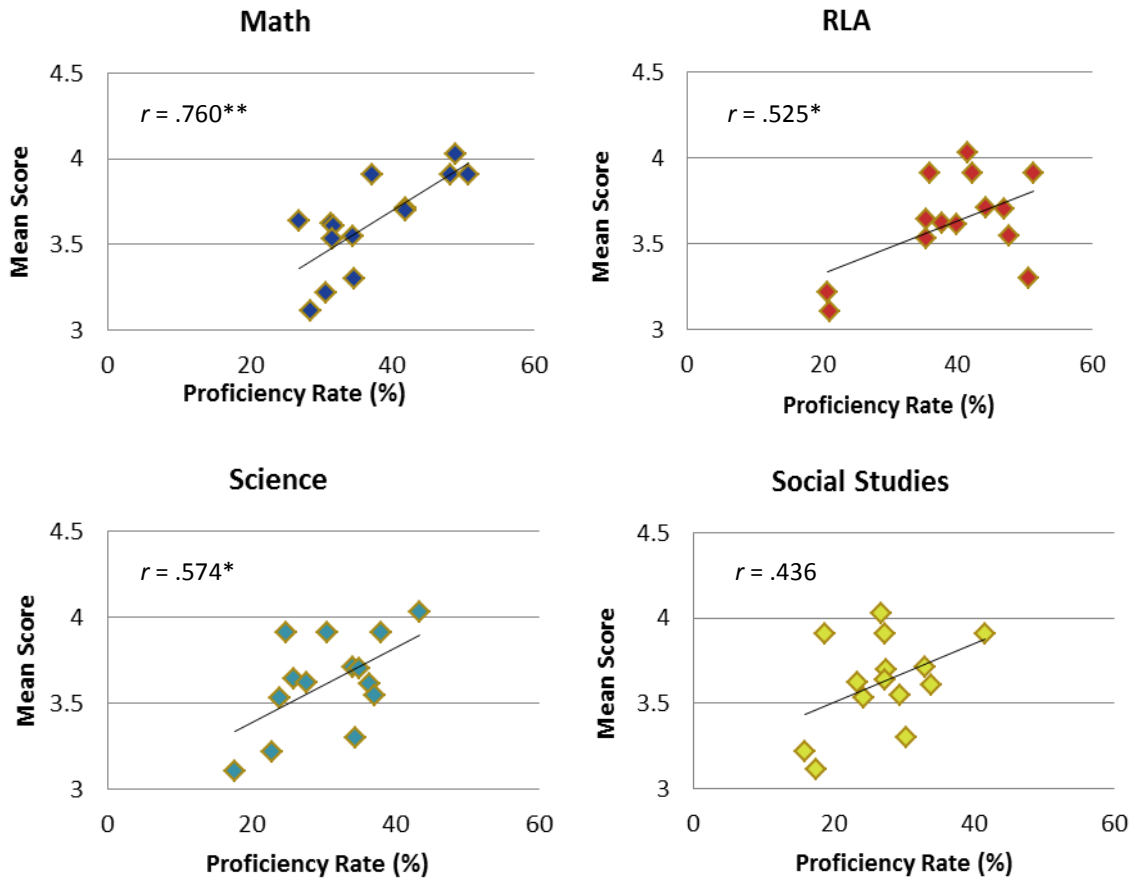


*Data from SY2009-2010. Data not available for SY2010-2011.

Correlation results between school culture survey and achievement data

We used Pearson's correlation coefficient (r) to measure the strength of relationship between school culture survey data and proficiency rates on the four content areas tested by the WESTEST 2 for SIG schools[10]. First, we tested the strength of relationship between school culture survey mean scores and proficiency rates on WESTEST 2 for the four content areas. Results showed that mean scores were positively correlated with proficiency rates for all content areas at least at the high end of a moderate relationship (Figure 63). Correlation with social studies proficiency rates was the weakest (r = .44) and was not statistically significant at the 0.05 level (Table 42). Proficiency rates for science and reading/language arts were strongly correlated, albeit at the low end, at the 0.57 and 0.53 levels respectively (Figure 63). Correlations for these two content areas were statistically significant at the 0.05 level. The strongest relationship was found between mathematics proficiency rates and culture survey mean scores (r = 0.76). This correlation was significant at the 0.01 level.

Figure 63. Scatter plots of 2010-2011 Proficiency Rates and School Culture Survey Mean Score



[10] *For help with interpreting Pearson's r, see the explanation provided in the Instructional Practices Inventory Results section (page 34).*

We also tested the strength of relationship between each of the seven domains that make up the school culture survey and proficiency rates on the four content areas. Results indicate that two domains, *Collaborative Leadership* and *Unity of Purpose*, were positively correlated with all content areas. With the exception of the correlation between social studies proficiency rates and Collaborative Leadership, all other correlations between these two domains and proficiency rates indicated strong positive relationships. Although the relationship between social studies proficiency rates and Collaborative Leadership is a moderate one (r = 0.46), it is worth pointing out that the strength of the relationship is on the high end of a moderate level. All correlation results between each of the domains and each of the content areas were significant at least at the 0.05 level (Table 42Table 42).

Two other domains, *Teacher Collaboration* and *Collegial Support*, showed a strong positive relationship with mathematics and science proficiency rates and correlations were statistically significant at the 0.05 level (Table 42). Although not statistically significant, results for correlations between Teacher Collaboration and proficiency rates for reading/ language arts and social studies indicate a moderate positive relationship (r = 0.44 and r = 0.34, respectively). Collegial Support was found to have a moderate positive relationship with reading/language arts (r = 0.32) and a slightly weaker positive relationship with social science (r = 0.25), but was not statistically significant in either case.

The *Learning Partnership* domain was positively correlated, at least at the moderate level, with proficiency rates for all four content areas (Table 42). Correlation was only statistically significant for math proficiency rates (p = 0.01). The *Professional Development* domain showed a strong positive relationship with math proficiency rates (r = 0.66) and the result was statistically significant at the 0.01 level. Correlations with the Professional Development domain and the other three content areas indicate relationships are comparatively much weaker, with only reading/language arts proficiency rates showing a low moderate positive relationship (r = 0.32). The last domain, *Efficacy Factor*, showed the weakest relationship with proficiency rates on the four content areas tested by the WESTEST 2. None of the correlation results tested statistically significant, although results indicate a moderate positive relationship with mathematics and science proficiency rates (r = 0.40 and r = 0.38, respectively) (Table 42).

Table 42. Bivariate Correlation Between Proficiency Rates and School Culture Survey Domains

	Correlation with content area proficiency rate			
	Mathematics	Reading/ language arts	Science	Social studies
Mean	.760**	.528*	.574*	.436
Collaborative Leadership	.630**	.648**	.535*	.459*
Teacher Collaboration	.788**	.443	.541*	.344
Professional Development	.658**	.255	.366	.262
Unity of Purpose	.742**	.599*	.636**	.536*
Collegial Support	.606*	.319	.512*	.246
Learning Partnership	.525*	.426	.348	.402
Efficacy Factor	.403	.221	.380	.201

*Correlation is significant at the 0.05 level (1-tailed).
**Correlation is significant at the 0.01 level (1-tailed).

Graduation and attendance rates

To provide additional evidence in response to EQ4, we calculated graduation and dropout rates for high schools and student attendance rates for all SIG schools, and will document the trend over the course of the program. Data from the 2010-2011 school year will serve as baseline data for subsequent reports.

Discussion

Evaluation Question 1

EQ1. To what extent has the SEA provided adequate support to LEAs and participating schools to successfully lead school improvement efforts?

Discussion

The document review revealed the following:

- *Professional development*. The West Virginia Department of Education (WVDE) provided four professional development sessions, attended by 425 participants from SIG schools and associated LEAs, which covered school improvement topics, including roles and responsibilities of school improvement specialists, school culture and data utilization, collaborative teams and curriculum, and collaborative teams and student leadership.
- *Technical assistance*. Additionally, WVDE Title I coordinators and school improvement specialists posted 358 technical assistance reports (for all 14 SIG schools), the vast majority of which involved assistance highly related to the topics covered in the professional develop—that is, fostering collaboration, strategic planning, and the collection and utilization of various types of data (e.g., Acuity, Instructional Practices Inventory, School Culture Survey) to drive school improvement efforts.

Based on a review of these topics, the professional development and technical assistance appeared well coordinated and focused on a cohesive set of school improvement objectives.

The following findings from the end-of-year survey, are based on a small group of respondents (n = 8), which represented district staff involved in all but one of the counties that have SIG schools.

- *Professional development*. Overall, respondents had favorable responses regarding the quality, usefulness, and relevance of the four professional development sessions and training materials provided during the sessions. Presenters were also well received in regards to their knowledge of the subject matter, clarity and effectiveness, as well as their helpfulness in answering questions during the sessions. Based on a retrospective pretest/posttest respondents indicated positive changes in their own knowledge, behaviors, and beliefs/attitudes as a result of the PD sessions they attended. Comments indicate that sessions were helpful in providing the most current best practices in regards to school improvement efforts and connecting topics/ themes covered at earlier professional development sessions.
- *Statewide system of support (SSOS) teleconferences*. Respondents indicated that these teleconferences were well organized, relevant, served as avenues of beneficial feedback, and helped improve communication among the different levels of their school system. Most respondents, however, regarded teleconferences as too time

consuming and a couple of respondents commented on the fact that topics were more aligned with the interest of the state department rather than their own work.

- *Technical assistance provided by the Office of Title I* to SIG schools was very highly received by respondents. Technical assistance providers were perceived to be very knowledgeable, committed, accessible, and extremely helpful to the school improvement process and comments indicate providers were instrumental in success gained by SIG schools thus far, particularly because technical assistance providers can attend to the needs of each SIG school directly.
- *Technical assistance provided by districts.* Responses indicate that SIG schools have benefited from the various types of services provided by districts: (a) transformation specialist; (b) professional development supported by SIG funds; (c) technical assistance related specifically to implementation of the SIG; and (d) monitored implementation of the SIG. Similar to services provided by the Office of Title I, comments indicate that this is particularly the case because services provided by districts can be individualized to schools in their districts.

  Challenges to school improvement efforts identified by respondents include issues around the amount of time away from schools, defining the roles of various stakeholders, and the unwillingness of some staff to change.

Limitations

Survey data are perceptual in nature and analysis and interpretations are descriptive in nature. Therefore, the reader should use some caution when reading the interpretations of the survey provided.

Evaluation Question 2

EQ2. To what extent has the SIG program resulted in improved quality and quantity of instruction and nonacademic supports in participating schools?

Discussion

Instructional Practices Inventory

The Instructional Practices Inventory (IPI) onsite data collection at the 14 SIG schools yielded the following results:

- An analysis of Pearson's correlation coefficient (r) between SIG IPI data (individual and various groupings of IPI categories) and proficiency rates on WESTEST 2 showed some statistically significant relationships:
  - *Individual IPI coding categories*: Teacher-led instruction (Category 4) was found to be positively correlated with reading/language arts proficiency rates. Student work with teacher engaged (Category 3) was negatively correlated with proficiency rates for reading/language arts and science while student work with teacher not engaged (Category 2) was negatively correlated with proficiency rates for mathematics.

- *Groupings of IPI coding categories*: The broad category of disengagement (Categories 1 and 2) was found to be negatively correlated with mathematics proficiency rates. The combination of student work with teacher not engaged and student work with teacher engaged (Categories 2 and 3) was negatively correlated with proficiency rates for reading/language arts and science. The grouping of IPI Categories 4, 5 and 6 was positively correlated with mathematics, reading/language arts, and science proficiency rates.

- The comparison of baseline IPI data for SIG schools to a typical school profile and comparison of SIG middle schools to historically successful and unsuccessful middle schools showed the following results:
  - Overall, and at each of the three programmatic levels, SIG schools have similar or slightly higher percentages of less desirable IPI categories and conversely, similar or slightly lower percentages of desirable IPI categories. The comparison at the group level, however, conceals considerable variation among individual SIG schools, a few of which have higher percentages of desirable categories and similar or lower percentages of less desirable categories.
  - In general, SIG middle schools showed slightly higher percentages of desirable categories and similar or lower percentages of less desirable categories compared to historically unsuccessful schools. Even though the ratio of the sum of percentages for Categories 4, 5, and 6 combined to Categories 1, 2, and 3 for SIG middle schools as a group is only slightly better than unsuccessful schools (1:1), three of the eight SIG middle schools showed a ratio (2:3) that is trending toward historically successful schools (1:3).

Effective School Practice Survey Results

Results of the Effective School Practice Survey ratings (for 11 of the 14 SIG schools) indicate that none of the mean domain scores reached 3.0 on the 4-point scale, which would have indicated strong fidelity to practices research has shown to be present in effective schools. This indicates that none of the domains are areas of particular strength for SIG schools as a whole at this point of implementation. However, SIG schools do fare better on a couple of domains compared with others and those domains can tentatively be considered areas of relative strength. With that in mind, the brief discussion below highlights what the survey data indicate to be areas of strength and weakness for SIG schools.

- The survey identified three areas of need for SIG schools as a group: *Parents and Learning*; *Curriculum, Assessment, and Instructional Planning*; and *Classroom Instruction*. In these domains, 30% or less had scores that met the threshold for above average fidelity (Table 43).

- The survey identified two areas of relative strength for SIG schools as a group: *Leadership and Decision Making* and *Professional Development*. In these domains, 45.5% or more SIG schools had scores that fell above 3.0, the threshold for above average fidelity (Table 43). However, at least 45% of SIG schools had scores below 3.0 for these domains, as well.

Table 43. Frequency of SIG Schools with above average fidelity score by domain

	SIG Schools	
Domain	n	%
Leadership and decision making	6	54.5
Professional development	5	45.5
Parents and learning	3	30.0
Curriculum, assessment, and instructional planning	3	27.3
Classroom instruction	3	27.3

- Analysis of Pearson's correlation coefficient (*r*) between Effective School Practice Survey data (overall mean and individual domain mean scores) and proficiency rates on WESTEST 2 showed some statistically significant relationships:
  - *Overall mean scores*. Results showed that overall effective school practice survey mean scores were positively correlated with proficiency rates for all content areas at least at the high end of a moderate relationship. With the exception of social studies, proficiency rates, results for all other content areas were statistically significant at least at the 0.05 level.
  - *Individual domain mean scores*. Three domains showed very strong positive relationships with proficiency rates and tested statistically significant at least at the 0.05 level. These were *Parents and Learning*, *Curriculum, Assessment, and Instructional Planning*, and *Classroom Instruction*. Of the three domains, Curriculum, Assessment, and Instructional Planning appeared to have the strongest correlation with proficiency rates, which tested statistically significant at the 0.01 level with proficiency rates for mathematics, reading/language arts, and science. Although the *Leadership* and *Professional Development* domains showed positive relationships with proficiency rates for the four content areas, only one out of eight correlations were statistically significant.

Limitations

There are several limitations of IPI data and analyses. First, the sample size of SIG schools (n = 14) is too small for methods of inferential statistics. What we presented above is descriptive in nature and findings should be taken with some caution. Second, *typical* IPI profile data are derived by averaging data submitted by schools that have used the IPI in the past. These schools are not selected systematically and/or randomly, which is to say that typical data are not *norms*; for this reason they cannot be considered to be statistically representative of a population. Third, although we would have preferred to use the average percentages of each category of the IPI data collected multiple times throughout the year, as mentioned earlier, we opted to limit our analysis to data that were available for all SIG schools. Fourth, IPI data for noncore classes were not available for four of the 14 schools and, as a result, we opted to focus our analyses primarily on core-class IPI data (mathematics, reading/language arts, science, and social studies) reported for SIG schools.

Regarding the Effective School Practice Survey, we remind the reader that findings are based on perceptual data and are descriptive in nature. While perceptual surveys can be great tools through which respondents can provide information about their feelings and

opinions about a specific topic/issue, they, however, are inherently biased. Therefore, the reader should use some caution when considering the interpretations of the Effective School Practice Survey provided above.

Evaluation Question 3

EQ3. To what extent has the SIG program resulted in improved academic achievement among students in participating schools?

Discussion

We conducted two analyses—one using student-level assessment data, which examined effects of the SIG program on academic achievement statewide; and another conducted at the school level, which provided feedback to each individual SIG school. Both analyses were conducted using the same sample of SIG and non-SIG comparison schools and WEST-EST 2 scores for mathematics and reading/language arts.

Program-level results

Overall, our results indicated that the matching process was relatively successful except with respect to Grade 8 in mathematics and Grades 7–9 in reading/language arts. Subsequent RM ANOVAs illustrated that, as far as the first SIG implementation year is concerned, there were relatively few significant differences between students in Cohort 1 SIG and non-SIG schools. However, some promising preliminary findings did emerge. For example, we found that the students in Cohort 1 SIG schools outperformed their counterparts in non-SIG comparison schools by a statistically significant margin in the Grade 3–4 cohort in both reading/language arts and mathematics, and also in the Grade 5–6 cohort in reading/language arts.

However, our enthusiasm about these findings must be tempered by the fact that we also found that students in the non-SIG comparison schools outperformed their counterparts in Cohort 1 SIG schools by a statistically significant margin in the Grade 6–7 cohort in reading/language arts. Table 44 provides a quick summary of where statistically significant interactions were found.

Table 44. Summary of Significant Program-Level Interaction Effects

Cohort	Content area	*p*	Group exhibiting higher achievement
Grade 3–4	Reading/language arts	.021	SIG
Grade 3–4	Mathematics	.021	SIG
Grade5–6	Reading/language arts	.022	SIG
Grade6–7	Reading/language arts	.022	Non-SIG Comparison

School-level results

Both the SIG and non-SIG comparison schools have illustrated some extraordinary gains during the past 3 years (i.e., SY2009–SY2011). However, in both reading/language arts and mathematics, more SIG schools increased their proficiency rates over time in both reading/language arts and mathematics than non-SIG comparison schools. Furthermore, as

a group, the Cohort 1 SIG schools had a median gain in reading/language arts that was 6.12% higher than the median gain for non-SIG comparison schools. For the same time period in mathematics, the SIG schools had a median gain that was 3.67% higher than the median gain for non-SIG comparison schools.

Gains have also occurred for both groups during the span between SY2010 and SY2011 which represents the first year of SIG implementation. Our results illustrate that in reading/language arts, the same number of schools in both groups (SIG and non-SIG comparison) improved their proficiency rates during this time. Furthermore, the median reading/language arts proficiency gain for Cohort 1 SIG schools during this period was very similar to non-SIG comparison schools (a difference of only 0.4%). However, with respect to mathematics, we found that approximately 78% of SIG schools increased their proficiency rates over the SY2010–SY2011 period while only 50% of non-SIG comparison schools were able to do so. The Cohort 1 SIG schools had a median gain of 5.4% in mathematics while the non-SIG comparison schools exhibited a median gain that was less than 1% (a difference of 5.3%).

Limitations

Several limitations of the program-level analyses must be considered. First, and most obvious, our failure to match groups adequately in Grade 8 mathematics and Grades 7–9 reading/language arts severely limits our view of the SIG program's impact in middle schools. This is especially problematic given that many of the Cohort 1 schools are middle schools. Another confounding factor arises from our decision to use panel data, that is, to follow individual students over time. While this allowed for higher precision in results because the within-student variability is lower, it also necessitated that we remove several students from the model (i.e., those that did not have two consecutive test scores at consecutive grade levels). In so doing, we have removed a significant number of students from these schools, and examined outcomes for a sample of students. We must ask if doing so gives us an entirely clear picture of school-wide achievement.

Second, and of critical importance, there were 14 different individual implementations of the SIG program during the first implementation year. Early on we understood that this would pose a massive resource problem if we attempted to study each one independently. Therefore, we made a conscious decision in consultation with Title I program staff that, given our limited resources and the fact that our audience consisted primarily of state-level administrators, we would design an evaluation that examined the impact of the SIG program as a whole, by focusing upon the elements that are common across implementations. This decision, while necessary, has implications in that the program-level analyses we conducted do not allow us to make any conclusions about the effectiveness of individual school implementations of the SIG that we know must vary quite widely. Still, we believe these analyses do address an important goal—to examine whether or not, as a whole, the Cohort 1 SIG schools are making improvements that are above and beyond what is occurring in similarly low performing schools that are not seeing such a massive influx of resources and substantial reforms.

We must also mention that our ultimate goal is to examine differences in student achievement trends over the entire course of the SIG program. However, this year's analysis was limited to include only one year of achievement data that has been collected post-SIG implementation. It is likely that the first implementation year of the SIG program has been used by schools primarily to start up many of the necessary reforms. As such, we must consider whether or not it is appropriate to expect to see many statistically significant differences in student achievement so early on. However, our results do at least provide some positive evidence that progress is occurring in SIG schools, even at this early stage. At the conclusion of SY2012, we will have a set of data with all SIG implementation data points, which will allow us to know more about how achievement has changed in these schools over the course of the program.

A final detail that bears mentioning is the fact that statistical power was low in several of our analyses (see Appendix C, page 131). This is important because it may lead to incorrectly attributing an effect upon student achievement to group membership when differences may actually be due to chance. Conversely, we may not have large enough sample sizes to detect very small, but statistically significant differences between groups. The latter is particularly problematic given that we observed very small effect sizes even when interactions were statistically significant. Further, because we had relatively small sample sizes within some grade level cohorts, we were unable to examine differences in achievement for various subgroups of students.

Regarding school-level results, the achievement profiles we have provided in this report are not statistically rigorous comparisons. For this reason, we do not draw conclusive inferences about the differences observed between SIG schools and non-SIG comparison schools, nor do we recommend readers do so. As noted previously, our intention in creating these profiles was to provide a point of reference for the types of gains these schools are experiencing. We hope the achievement profiles will be useful to SIG schools and to those providers assisting SIG schools when used in conjunction with additional data sources.

Readers should also keep in mind that the school matching process was not perfect. As noted in the Results section, we considered several variables using different tolerances to match our schools. However, we always prioritized the match based upon reading/language arts and mathematics proficiency. We were not able to control for other unmeasured differences between the schools. Nor were we able to conduct a statistical examination of the extent to which the matching process at the school level was successful. Further, we must reiterate that we matched schools on achievement, based upon their SY2010 proficiency rates since this represented the most recently available pre-SIG implementation point. The fact that we did not match upon SY2009 data is clear from many of the charts. This is somewhat problematic when examining trends from SY2009–SY2011, as we have not controlled for starting points. However, it is also clear that our efforts to match schools based upon their 2010 proficiency rates were somewhat more successful in most cases. For this reason, we put more faith in the differences observed among groups between SY2010 and SY2011.

The proficiency rates we calculated are not those that were used for adequate yearly progress determinations. Our rates do not include students who take the state's alternate assessment (Alternate Performance Task Assessment [APTA]). Further, we have included

testing information for all students regardless of whether or not the school is formally held accountable for these students. For example, in high schools, our proficiency rate data include students in Grades 9 through 11. Typically, secondary schools are only accountable for Grade 11 students. Further, our data may also include students who were not enrolled at the school for the full academic year. Also, we have constructed proficiency rates for SY2009 based upon the SY2010 operational cut scores in order to ensure accuracy of our trend information. Therefore, there will be a large discrepancy between the proficiency rates reported for AYP in 2009, and those reported here.

Finally, we must acknowledge that the analyses in our achievement profiles and our program-level analyses do not include data for the full SIG implementation period (i.e., SY2010–SY2012). As we have indicated previously, it is likely that the SIG program has not reached full implementation at this point in most schools. Therefore, we caution against making broad judgments about the program's effectiveness at this point. This is an interim report and should be treated as such. The information should be used to guide implementation efforts, not to make sweeping judgments about the program.

Evaluation Question 4

EQ4. To what extent have SIG schools successfully implemented and institutionalized practices and structures that are supportive of continuous school improvement?

Discussion

Measure of Schools Capacity for Improvement

The Measure of Schools Capacity for Improvement (MSCI), a survey that was taken by 571 faculty and staff members in the SIG schools, yielded the following results:

- The MSCI identified at least two critical areas of need for SIG schools as a group. These are Capacity for Peer Reviewed Practice and Increasing Expectations for Student Performance. These were areas where the median for SIG schools fell very near the 50th percentile.
- The MSCI identified three areas of strength for SIG schools as a group. These are Differentiated Instruction, Technical Resources, and Improvement Program Coherence. These are areas where the median for SIG schools fell near or above the 70th percentile. As such, these areas appear to be relatively well-covered in most SIG schools. There were, however, a few schools that rated quite low in these domains, and their performance will continue to need monitoring and intervention.

School Culture Survey

Results of the School Culture Survey identified at least three areas of need for SIG schools as a group: *Learning Partnerships*, *Teacher Collaboration*, and *Collaborative Leadership*. These domains had average scores for SIG schools that fell very near or below the overall mean for all seven dimensions.

The School Culture Survey also identified at least three areas of strength for SIG schools as a group: *Unity of Purpose*, *Collegial Support*, and *Professional Development*. These are areas where domain average scores for SIG schools fell above the overall mean for all seven dimensions. There were, however, a few schools that rated quite low in these domains, and their performance will continue to need monitoring and intervention.

Analysis of Pearson's correlation coefficient (r) between school culture survey data (overall mean and individual domain mean scores) and proficiency rates on WESTEST 2 showed some statistically significant relationships:

- *Overall mean scores*. Results showed that overall school culture mean scores were positively correlated with proficiency rates for all content areas at least at the high end of a moderate relationship. With the exception of social studies proficiency rates, results for all other content areas were statistically significant at least at the 0.05 level.
- *Individual domain mean scores*. The domains *Collaborative Leadership* and *Unity of Purpose* were found to have the strongest correlation with proficiency rates for all content areas. Correlations between these two domains and proficiency rates for all four content areas were statistically significant at least at the 0.05 level. *Teacher Collaboration* had a positive relationship with all four content area proficiency rates at least at the moderate level, although only two of the correlations were statistically significant. *Collegial Support* was correlated positively at the moderate level with proficiency rates for three out of the four content areas and two were statistically significant. Although none of the correlations between proficiency rates for the four content areas and the domain *Learning Partnership* was found to be statistically significant, the domain was positively correlated with all content area proficiency rates at least at the moderate level. Although the domains *Professional Development* and *Efficacy Factor* showed positive relationships with proficiency rates for the four content areas, only one out of eight correlations were statistically significant.

Results validate the value of the school culture variables to SIG schools. Proficiency rates for all content areas are highly correlated with overall mean scores on the survey.

Limitations

The MSCI findings are descriptive in nature and are based on perceptual data. While it is interesting and useful to compare the results of SIG schools with norm-referenced schools, there is not sufficient evidence to determine whether these differences between SIG schools and norm-referenced schools were meaningfully and statistically significant.

The School Culture Survey findings are, also, based on perceptual data and descriptive in nature. Therefore, the reader should use some caution when reading the interpretations of the School Culture Survey provided above.

Recommendations

Readers are cautioned to view these recommendations as based on only one year of implementation and data collection. In some cases (e.g., the Instructional Practices Index [IPI]), data were collected early in the year, and therefore, represent pre-SIG intervention, baseline observations. Keeping these limitations in mind, we offer recommendations for implementation going forward, and for ongoing monitoring and evaluation of the SIG program.

Implementation

Based on document reviews and responses to the various surveys conducted for this evaluation, there were some areas of practice that show movement in a positive direction, and we therefore encourage the Office of Title I to continue to—

- Link professional development and technical assistance to each other and to cohesive and well-defined goals for improvement, as was done during the 2010–2011 school year.
- Provide professional development that LEA Title I directors perceive as high quality, useful, and relevant, and as producing positive change in their own knowledge, behaviors, and beliefs and attitudes.
- Maintain the quality of support provided by districts in 2010-2011. Evidence from this study suggests that, when compared to professional development, district-provided support affords greater opportunity to differentiate intervention activities based on specific school needs.

Other evidence from this study suggest where efforts need to be intensified, continued, or adjusted:

- Use SSOS teleconferences, as much as possible, as two-way conversations to elicit information about participants' *current work* and to provide feedback that would help guide their school improvement efforts.
- Use more onsite technical assistance, as well as district level services, to help alleviate one of the most recurring concerns expressed by SIG school staff—that is, time spent away from their school buildings participating SIG activities.
- Provide additional guidance to elucidate the roles and responsibilities of all those involved in the school improvement effort in SIG schools and their home districts.
- Provide guidance on appropriate ways to handle staff members who are resistant to efforts to bring about positive changes in SIG schools.
- Encourage all SIG schools to collect IPI data multiple times per year.
- Provide guidance on ways schools can increase student in higher order/deeper learning (IPI Categories 5 and 6), and reduce the incidence of student work with teacher not engaged (IPI Category 2) and student work with teacher engaged (IPI Category 3). Table 16 (page 41) provides detailed information with school scores that fell below the *favorable* range.Focus additional professional development, technical

assistance, and other resources on addressing effective school practices, which evidence from the Effective School Practice Survey suggests need more attention. Domains that seem to need the most attention by the greatest number of schools include *Parents and Learning*; *Curriculum, Assessment, and Instructional Planning*; and *Classroom Instruction*. These domains not only had the lowest mean scores, but were also the most strongly correlated with student achievement data on WESTEST 2, suggesting that improvement in these areas may have a tangible impact on student achievement.Table B- 3 through Table B- 7 in Appendix B, beginning on page 126, provide detailed information about each school, with school scores that fell below favorable ratings highlighted.

- Focus additional professional development, technical assistance, and other resources to address deficiencies in SIG schools related, especially, to the Measure of School's Capacity for Improvement (MSCI) subscales *Peer Reviewed Practice* (i.e., the observation and review by staff of their peers' work), and *Expectations for Student Performance* (i.e., staff members' expectations of the students and their beliefs that all students can perform well academically). Figure 48 through Figure 54 (beginning on page 82) display results for all seven domains by school.
- Focus additional professional development, technical assistance, and other resources on addressing widespread deficiencies in SIG schools in the School Culture Survey domains of *Learning Partnerships* (i.e., the degree to which teachers, parents and students work together for the common good of the student), *Teacher Collaboration* (i.e., the degree to which teachers engage in constructive dialogue that furthers the educational vision of the school), and *Collaborative Leadership* (i.e., the degree to which school leaders establish and maintain collaborative relationships with school staff). Two of these domains, *Teacher Collaboration* and *Collaborative Leadership*, not only had the lowest mean scores, but were also the most strongly correlated with student achievement data on WESTEST 2, suggesting that improvement in these areas may have a tangible impact on student achievement. Table E- 1 through Table E-7 in Appendix E, beginning on page 154, provide detailed information about each school, with school scores that fell below favorable ratings highlighted.

Ongoing Monitoring and Evaluation

- Reserve holistic judgments of the SIG program's impact upon student achievement until achievement data spanning all years of the SIG program are collected and analyzed thoroughly. Our interim evaluation of student achievement in SIG schools should be used diagnostically as a "temperature check" of the SIG program and to identify schools where relatively minor gains are occurring when compared with other SIG schools and matched non-SIG schools. Using these data in this manner allows WVDE to intervene in these sites before an opportunity for course correction—using the substantial resources provided by the SIG—is lost.
- Identify a more robust matching methodology for our final round of program-level analyses, one that allows for adequate student-level matching at all programmatic levels. Accomplishing this degree of a match will be absolutely critical if we wish to

provide a fair and accurate view of the SIG program. The Office of Research in consultation with Title I staff must also consider whether a longitudinal panel data approach is the most appropriate method to evaluate the SIG program given the level of attrition that occurs with student records.

- Investigate whether or not the program level effects on WESTEST 2 mathematics and reading/language arts scores found in elementary schools reflect some set of innovative practices. If so, these practices should be identified, nurtured, and, if possible, replicated in middle/high schools.
- Continue to encourage SIG schools to administer the School Culture Survey on an annual basis and utilize the data for school improvement efforts.
- Utilize average percentages of IPI data collected at least two times during each school year in future evaluations.
- In addition to LEA Title I coordinators, extend the invitation to participate in the 2011-2012 End-of-Year Survey to school leadership team members and school improvement/transformation specialists, who can contribute valuable feedback regarding the quality, relevance, and usefulness of professional development and technical assistance schools and districts receive during the school year.
- Encourage all school improvement specialists to participate in the Effective School Practice Survey. It is clear from the frequency of "I don't know" responses from SEA Title I directors that school improvement specialists are best situated to provide relevant feedback on many of the survey items included in the survey. Re-examine the correlation between the Effective School Practice Survey and WESTEST 2 proficiency once this wider audience has participated.
- Eliminate MSCI survey from analysis at least until the final year of SIG program for the following reasons: (a) there is a tremendous amount of overlap among the MSCI, School Culture, and Effective School Practice surveys; (b) the MSCI survey is best used as a diagnostic instrument at the onset of program implementation to gauge schools' capacity for improvement and perhaps at the end of a program cycle to measure any gains obtained as a result of improvement efforts; and (c) eliminating the MSCI survey will reduce the burden of completing an additional survey by school staff in SIG schools who are already tasked with a great many responsibilities.

References Cited

Gruenert, S. & Valentine, J. (1998). *School culture survey*. Columbia, MO: Middle Level Leadership Center, University of Missouri.

Hughes, G. K., Copley, L. D., Howley, C. W., & Meehan, M. L. (2006). *Measure of School Capacity for Improvement (MSCI) user manual and technical report*. Charleston, WV: Edvantia, Inc.

Perlman, C.L., & Redding, S. (Eds.) (2009). *Handbook on effective implementation of School Improvement Grants*. Lincoln, IL: Center on Innovation & Improvement. Retrieved from http://www.centerii.org.

Redding, S. & Laboratory for Student Success (2006). *The mega system. Deciding. Learning. Connecting. A handbook for continuous improvement within a community of the school.* Lincoln, IL: Academic Development Institute. Retrieved from http://www.adi.org/mega/chapters/ChapterOne.pdfHughes, et al.

U.S. Department of Education (2010a). *Guidance on School Improvement Grants under Section 1003(g) of the Elementary and Secondary Education Act of 1965.* Revised June 29, 2010. Washington, DC: Author, Office of Elementary and Secondary Education. Retrieved from http://www2.ed.gov/programs/sif/guidance20100120.doc.

U.S. Department of Education. (2010b). *School Improvement Grants application: Section 1003(g) of the Elementary and Secondary Education Act. Fiscal year 2010.* CFDA Number 84.377A. Washington, DC: Author, Office of Elementary and Secondary Education. Retrieved from http://www2.ed.gov/programs/sif/applicant.html.

Valentine, Jerry (2005). *The Instructional Practices Inventory: A process for profiling student Engaged learning for school improvement*. Columbia, MO: University of Missouri, Middle Level Leadership Center.

Valentine, Jerry (2009a). *IPI profiles: Baseline and subsequent observation data*. Columbia, MO: University of Missouri, Middle Level Leadership Center.

Valentine, Jerry (2009b). *IPI baseline and all profile averages by school grade levels*. Columbia, MO: University of Missouri, Middle Level Leadership Center.

Valentine, Jerry (2011). *IPI research summary A: IPI data and student achievement on high stakes tests*. University of Missouri, Columbia, MO: Middle Level Leadership Center.

Walberg, H. J. (Ed.) (2007). *Handbook on restructuring and substantial school improvement.* Lincoln, IL: Center on Innovation & Improvement. Retrieved from http://www.centerii.org/survey/.

West Virginia Department of Education (2010a). *State of West Virginia consolidated state application accountability workbook for state grants under Title IX, Part C, Section 9302 of the Elementary and Secondary Education Act (Public Law 107-110).* Revised June 2010. Charleston, WV: Author. Retrieved from http://oepa.state.wv.us/.

West Virginia Department of Education (2010b). *Title I 1003(g) School Improvement Grants*. PowerPoint presentation, March 9, 2010. Charleston, WV: Author, Office of Title I. Retrieved from http://wvde.state.wv.us/titlei/documents/1003gSIGPresentation-March2010_000.ppt.

West Virginia Department of Education (2010c). *West Virginia first state to receive federal turnaround funds*. Press release, March 11, 2010. Charleston, WV: Author. Retrieved from http://wvde.state.wv.us/news/2051/.

West Virginia Department of Education (2010d). *West Virginia funded districts summary: School Improvement Grant 2010-2011, Tier I and Tier II schools, Cohort I*. Charleston, WV: Author. Retrieved from http://wvde.state.wv.us/titlei/documents/WVSIGDistrictFundingSummary.pdf.

West Virginia Department of Education (n.d.) *West Virginia definition of persistently lowest achieving schools*. Charleston, WV: Author. Retrieved from http://wvde.state.wv.us/titlei/1003eligible.html.

West Virginia Department of Education. (2010e). *West Virginia School Improvement Grant application. Fiscal year 2010*. Charleston, WV: Author, Office of Title I.

Appendix A. End-of-Year Survey Tables

Table A- 1. End-of-Year Survey Respondents' Perception of the Quality of Professional Development

Survey Item	Response counts and percent*											
	Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
	N	%	N	%	N	%	N	%	N	%	Mean	SD
The professional development I attended was…												
…high quality.	0	0.0	0	0.0	2	25.0	2	25.0	4	50.0	4.25	0.21
…well organized.	0	0.0	0	0.0	0	0.0	4	50.0	4	50.0	4.5	0.27
…intensive in nature.	0	0.0	0	0.0	1	12.5	4	50.0	3	37.5	4.25	0.23
…specific and content-focused.	0	0.0	0	0.0	1	12.5	3	37.5	4	50.0	4.38	0.23
…relevant to my current needs and circumstances as an educator.	0	0.0	1	14.3	0	0.0	2	28.6	4	57.1	4.29 **	0.24
…hands-on and included active learning opportunities.	0	0.0	1	12.5	1	12.5	3	37.5	3	37.5	4	0.17
…supported by follow-up discussion or collaboration at our school or office or online.	0	0.0	0	0.0	1	12.5	3	37.5	4	50.0	4.38	0.23
…supported by related follow-up PD sessions.	0	0.0	0	0.0	2	25.0	2	25.0	4	50.0	4.25	0.21
…beneficial and had a positive impact on our students and/or school, overall.	0	0.0	0	0.0	2	25.0	1	12.5	5	62.5	4.38	0.26
Overall Quality Mean											4.3	0.14

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree
**Only 7 respondents answered this survey item

Table A- 2. End-of-Year Survey Respondents' Perception of the Quality of Training Materials/Resources

Survey Item	Response counts and percent*											
	Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
	N	%	N	%	N	%	N	%	N	%	Mean	SD
Professional development training materials/resources were…												
…provided.	0	0.0	0	0.0	0	0.0	3	37.5	5	62.5	4.63	0.29
…relevant to my work.	0	0.0	1	12.5	0	0.0	3	37.5	4	50	4.25	0.23
…adequate.	0	0.0	0	0.0	1	12.5	4	50	3	37.5	4.25	0.23
…high quality.	0	0.0	0	0.0	1	12.5	4	50	3	37.5	4.25	0.23
…useful to my work.	0	0.0	1	12.5	0	0.0	3	37.5	4	50	4.25	0.23
Overall Quality											4.33	0.17

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree

Table A- 3. End-of-Year Survey Respondents' Perceptions of the Quality of Presenters

Survey Item	Response counts and percent*											
	Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
	N	%	N	%	N	%	N	%	N	%	Mean	SD
Presenters at the professional development...												
…were knowledgeable about the topic.	0	0.0	0	0.0	0	0.0	5	62.5	3	37.5	4.38	0.29
…were well organized.	0	0.0	0	0.0	2	25.0	3	37.5	3	37.5	4.13	0.19
…presented the material clearly and effectively.	0	0.0	0	0.0	0	0.0	5	62.5	3	37.5	4.38	0.29
…answered questions raised during sessions adequately.	0	0.0	0	0.0	0	0.0	5	62.5	3	37.5	4.38	0.29
Overall Quality											4.32	0.12

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree

Table A- 4. End-of-Year Survey Respondents' Perceptions of the Impact of the Professional Development Sessions

	Response counts and percent											
	Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
Component	N	%	N	%	N	%	N	%	N	%	Mean	SD
I was knowledgeable about the topics covered in the school improvement professional development.												
Before the professional development	0	0.0	0	0.0	2	25.0	6	75.0	0	0.0	3.75	0.33
After the professional development	0	0.0	0	0.0	0	0.0	4	50.0	4	50.0	4.50	0.27
Mean Difference (Gain)											0.75	
I practiced the behaviors and/or skills that were taught in the school improvement professional development.												
Before the professional development	0	0.0	0	0.0	1	12.5	7	87.5	0	0.0	3.88	0.38
After the professional development	0	0.0	0	0.0	0	0.0	5	62.5	3	37.5	4.38	0.29
Mean Difference (Gain)											0.50	
I held the attitudes/beliefs that the school improvement professional development encouraged												
Before the professional development	0	0.0	0	0.0	1	12.5	6	75.0	1	12.5	4.00	0.31
After the professional development	0	0.0	0	0.0	1	12.5	2	25.0	5	62.5	4.50	0.26
Mean Difference (Gain)											0.50	

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree

Table A- 5. End-of-Year Survey Respondents' Perceptions of the Quality of SSOS Teleconferences

	Response counts and percent											
	Strongly disagree		Disagree		Agree		Strongly agree		Not applicable			
Survey Item	N	%	N	%	N	%	N	%	N	%	Mean*	SD
The statewide system of support (SSOS) teleconferences (Collaborative Conversations)…												
…are well organized.	0	0.0	1	12.5	4	50.0	3	37.5	0	0.0	3.25	0.23
…are relevant to my work.	0	0.0	1	12.5	4	50.0	3	37.5	0	0.0	3.25	0.23
…are too time consuming.	0	0.0	5	62.5	1	12.5	2	25.0	0	0.0	2.38**	0.27
…have been avenues of beneficial feedback.	0	0.0	2	25.0	4	50.0	2	25.0	0	0.0	3.00	0.20
…have helped to improve communication between the different levels of my school system.	0	0.0	3	37.5	2	25.0	3	37.5	0	0.0	3.00	0.18
…have helped to improve collaboration between the different levels of my school system.	0	0.0	3	37.5	3	37.5	2	25.0	0	0.0	2.88	0.18
…have helped me see how the various pieces of the SIG program fit together.	0	0.0	2	28.6	4	57.1	1	14.3	0	0.0	2.86***	0.24
Overall Quality											2.94	0.3

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree
**Scale: 1 = Strongly agree, 2 = Agree, 3 = Disagree, 4 = Strongly disagree
***Only 7 respondents answered this survey item

Table A- 6. End-of-Year Survey Respondents' Perception of the Quality of Technical Assistance

	Response counts and percent											
	Strongly disagree		Disagree		Agree		Strongly agree		Not applicable			
Survey Item	N	%	N	%	N	%	N	%	N	%	Mean*	SD
Please respond to the following statements about the technical assistance services your SIG schools received from the WVDE Office of Title I.												
The technical assistance was high quality.	1	12.5	0	0.0	3	37.5	4	50.0	0	0.0	3.25	0.23
The technical assistance was relevant to the SIG.	1	12.5	0	0.0	2	25.00	5	62.5	0	0.0	3.38	0.27
The technical assistance was useful.	1	12.5	0	0.0	2	25.00	5	62.5	0	0.0	3.38	0.27
The technical assistance helped us to implement school improvement efforts.	1	12.5	0	0.0	2	25.00	5	62.5	0	0.0	3.38	0.27
Overall Quality											3.35	0.06

*Scale: 1 = Strongly agree, 2 = Agree, 3 = Disagree, 4 = Strongly disagree

Table A- 7. End-of-Year Survey Respondents' Perceptions of the Quality of Technical Assistance Provider

	Response counts and percent											
	Strongly disagree		Disagree		Agree		Strongly agree		Not applicable			
Survey Item	N	%	N	%	N	%	N	%	N	%	Mean*	SD
Please respond to the following statements about the technical assistance provider.												
The provider was well prepared and organized.	1	12.5	0	0.0	2	25.0	5	62.5	0	0.0	3.38	0.27
The provider was knowledge-able about the topic.	1	12.5	0	0.0	2	25.0	5	62.5	0	0.0	3.38	0.27
The provider was helpful.	1	12.5	0	0.0	1	12.5	6	75.0	0	0.0	3.50	0.34
Overall Quality											3.42	0.07

*Scale: 1 = Strongly agree, 2 = Agree, 3 = Disagree, 4 = Strongly disagree

Table A- 8. End-of-Year Survey Respondents' Perceptions of the Benefit to SIG Schools from Services Provided by Districts

	Response counts and percent											
	Strongly disagree		Disagree		Agree		Strongly agree		Not applicable			
Survey Item	N	%	N	%	N	%	N	%	N	%	Mean*	SD
Please indicate the extent to which you agree or disagree with the following statements.												
SIG schools benefited from the transformation specialist hired by the district.	1	12.5	0	0.0	2	25.0	5	62.5	0	0.0	3.38	0.27
SIG schools benefited from professional development provided by the district.	1	12.5	0	0.0	3	37.5	4	50.0	0	0.0	3.25	0.23
SIG schools benefited from technical assistance related to the implementation of the SIG provided by the district.	1	12.5	0	0.0	3	37.5	4	50.0	0	0.0	3.25	0.23
SIG schools benefited from monitoring of the SIG provided by the district.	1	14.3	0	0.0	2	28.6	4	57.1	0	0.0	3.29**	0.24
Overall Quality											3.29	0.06

*Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Agree, 4 = Strongly agree
**Only 7 respondents answered this survey item

Appendix B. Effective School Practice Survey Tables

Table B- 1. Effective School Practice Survey Items by Domain

Item #	Item
	Leadership and Decision-Making (14 Items)
1	Along with the principal, which of the following options best describes the composition of the Leadership Team?
2	How often does the Leadership Team meet?
3	How would you characterize the level of communication between the Leadership Team and the faculty and staff?
4	Which teachers in this school participate in grade-level, grade-level cluster, or subject-area Collaborative Teams?
5	How frequently do most Collaborative Teams meet?
6	What is your estimate of how many staff have a good understanding of their roles and responsibilities as members of their Collaborative or Leadership Team?
7	How often does staff use DP21 data when discussing school strengths and weaknesses?
8	How frequently does the Leadership Team communicate DP21 data with teams and teachers?
9	How often are decisions about school improvement based on data?
10	To what degree do Leadership Team members monitor school-level student learning data to set yearly learning goals?
11	To what degree does the principal actively participate in the school's Collaborative Teams?
12	How often does the principal monitor curriculum and classroom instruction?
13	How often does the principal provide reports on the strengths and areas of needed improvement of the faculty as a whole, based on classroom observations utilizing indicators of effective teaching?
14	How often does the principal plan opportunities for teachers to share their strengths with other teachers through peer-to-peer collaboration or mentoring?
	Professional Development (3 Items)
1	How often is professional development provided based on aggregate needs detected during classroom observations utilizing indicators of effective teaching?
2	How often does teacher evaluation examine the same indicators used in planning for professional development?
3	How often does the principal conduct classroom observations to examine application of skills/ knowledge gained from professional development?
	Parents and Learning (6 Items)
1	How often do parents receive jargon-free communication about learning standards, their children's progress, and the parents' role in their children's school success?
2	How often do parents receive practical guidance to maintain daily conversations with their children about their school experiences and progress?
3	How often do parents receive practical guidance to encourage their children's regular reading habits at home?
4	How often do parents receive practical guidance on how to model and encourage respectful and responsible behaviors?
5	How often are parents provided opportunities to meet with each other to share their child-rearing concerns and successes?

Table B- 1 continued next page

Table B- 1. Effective School Practice Survey Items by Domain

Item #	Item
6	How often are parents provided opportunities to meet with teachers to discuss both their children's progress in school and their children's home-based study and reading habits?
	Curriculum, Assessment, and Instructional Planning (10 Items)
1	How often do units of instruction include standards-based objectives and criteria for mastery?
2	How often are objectives leveled to target learning to each student's demonstrated prior mastery, based on multiple points of data (e.g., unit tests and student work)?
3	How often do teachers individualize instruction based on pretest results to provide support for some students and enhanced learning opportunities for others?
4	How often do teachers reteach based on post-test results?
5	How often do Collaborative Teams develop materials for their standards-aligned learning activities and share the materials among themselves?
6	How often is data entry into WVEIS completed in a timely fashion?
7	How often are data validated after initial entry into WVEIS by end users (e.g., teachers, office staff)?
8	How often are WVEIS reports generated?
9	How often are WVEIS reports used to assess strengths and weaknesses of the curriculum and instructional strategies?
10	How often does school staff use student learning data from WVEIS to identify students in need of instructional support or enhancement, and plan accordingly?
	Classroom Instruction (15 Items)
1	How often do teachers develop weekly lesson plans based on aligned units of instruction with clearly stated lesson topics, themes, and learning objectives?
2	How often do teachers assess student progress using a variety of evaluation methods and maintain a record of the results?
3	How often do teachers differentiate assignments (individualize instruction) in response to individual student performance on pretests and other methods of assessment?
4	How often do teachers use a variety of instructional methods, including modeling, demonstration, and graphics, and transition effectively?
5	How often do teachers explain directly and thoroughly?
6	How often do teachers review with open-ended questioning and encourage elaboration?
7	How often do teachers summarize key concepts?
8	How often do teachers encourage or provide opportunities for peer interaction?
9	How often do teachers encourage students to paraphrase, summarize, and relate?
10	How often do teachers encourage students to check their own comprehension?
11	How often do teachers check, mark, and return homework with comments?
12	How often do teachers report to parents the student's mastery of specific standards-based objectives?
13	How often do teachers maintain well-organized student learning materials in the classroom?
14	How often do teachers reinforce classroom rules and procedures by positively teaching them?
15	How often do teachers engage all students?

Table B- 2. Distribution of "I don't know" Responses, by Raters, by Effective School Practice Survey Domain

Raters (# of schools rated)	Leadership and decision making		Professional develop-ment		Parents and learning		Curriculum, assessment, and instructional planning		Classroom instruction		Total # of survey items	
	N	%	N	%	N	%	N	%	N	%	N	%
SEA Title I coordinator (12)	13	7.7	11	30.6	53	73.6	54	45.0	28	15.6	159	15.1
LEA Title I coordinator (6)	0	0.0	0	0.0	6	16.7	10	16.7	9	10.0	25	2.4
School improvement specialist (4)	0	0.0	1	8.3	0	0.0	0	0.0	1	1.7	2	0.00
Total	13	4.2	12	18.2	59	44.7	64	29.1	38	11.5	186	17.6

Table B- 3. Average Effective School Practice Survey Subscale Scores for Leadership and Decision Making, by Item, by School

		Leadership and decision making (14 items)													
School (number of raters)	Mean	Composition of leadership team	Leadership team meeting frequency	Level of communication	Participation in collaborative teams	Collaborative team meeting frequency	Faculty level of understanding of roles and responsibilities	Use of DP21 data	DP21 data communication frequency	Data based decision making	Data based yearly learning goals	Principal participation in collaborative teams	Curriculum and classroom observation	Observation based effective teaching feedback	Peer-to-peer collaboration opportunities
School N (1)	3.07	4.00	4.00	4.00	4.00	4.00	3.00	3.00	3.00	4.00	4.00	**1.00**	**2.00**	**2.00**	**1.00**
School F (2)	2.68	**1.50**	3.00	3.00	3.50	4.00	3.00	**1.50**	**1.50**	3.50	3.50	2.50	3.00	**2.00**	**2.00**
School D (3)	3.36	3.33	3.00	3.33	4.00	4.00	3.67	3.00	3.33	4.00	3.67	3.33	3.00	2.67	2.67
School C (4)	3.16	3.50	4.00	3.25	4.00	4.00	**2.25**	**2.25**	**2.25**	3.50	3.25	3.50	3.50	3.00	**2.00**
School L (1)	**2.31**	**2.00**	**2.00**	**2.00**	4.00	4.00	**2.00**	**2.00**	**1.00**	**2.00**	3.00	**1.00**	4.00	ND	**1.00**
School K (2)	3.00	**1.50**	3.00	3.00	4.00	3.00	3.00	**2.00**	**1.50**	3.50	3.00	4.00	4.00	4.00	2.50
School I (3)	3.25	2.67	3.67	3.00	4.00	4.00	3.00	**2.33**	**2.33**	3.67	3.33	3.33	3.67	3.00	3.50
School M (2)	**2.14**	**1.50**	3.50	**2.00**	**2.00**	3.00	**1.50**	**2.00**	**2.00**	3.00	3.00	**2.00**	**1.50**	**1.50**	**1.50**
School A (2)	3.68	3.50	4.00	3.50	3.50	4.00	3.50	3.00	3.00	4.00	4.00	4.00	4.00	4.00	3.50
School H (1)	**2.22**	**1.00**	3.00	**2.00**	4.00	ND	ND	**2.00**	**1.00**	3.00	**2.00**	**2.00**	ND	ND	ND
School J (1)	2.85	**2.00**	**2.00**	**2.00**	4.00	4.00	**2.00**	**2.00**	**1.00**	4.00	4.00	4.00	4.00	ND	**2.00**
Mean	2.88	**2.41**	3.20	2.83	3.73	3.80	2.69	**2.28**	**1.99**	3.47	3.34	2.79	3.27	2.77	**2.17**
SD	0.50	1.03	0.73	0.71	0.61	0.42	0.71	0.51	0.86	0.61	0.60	1.15	0.90	0.92	0.89

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity
ND = No data
Red indicates the rounded mean rating falls below the threshold for above average fidelity.

Table B- 4. Average Effective School Practice Survey Subscale Scores for Professional Development, by Item, by School

		Professional development (3 items)		
School (number of raters)	Mean*	Data driven professional development	Teacher evaluation based on professional development	Observation based on professional development gains
School N (1)	**2.00**	3.00	ND	**1.00**
School F (2)	3.17	3.50	4.00	**2.00**
School D (3)	3.22	4.00	3.00	2.67
School C (4)	2.92	3.00	2.50	3.25
School L (1)	**2.00**	**1.00**	ND	3.00
School K (2)	3.00	3.00	**2.00**	4.00
School I (3)	2.61	3.33	**2.00**	2.50
School M (2)	**2.00**	2.50	**2.00**	**1.50**
School A (2)	3.50	3.50	3.50	3.50
School H (1)	**2.00**	**2.00**	ND	ND
School J (1)	4.00	4.00	ND	4.00
Mean	2.77	2.98	2.71	2.74
SD	0.70	0.89	0.81	1.01

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity
ND = No data
Red indicates the rounded mean rating falls below the threshold for above average fidelity.

Table B- 5. Average Effective School Practice Survey Subscale Scores for Parents and Learning, by Item, by School

		Parents and learning (6 items)					
School (number of raters)	Mean*	Jargon-free communication	Guidance for daily conversation	Guidance to encourage reading habits	Guidance to encourage good behaviors	Parental meeting opportunities	Parent-teacher meeting opportunities
School N (1)	**2.00**	**2.00**	ND	ND	ND	ND	ND
School F (2)	**1.33**	**2.00**	**1.00**	**2.00**	**1.00**	**1.00**	**1.00**
School D (3)	**2.42**	3.50	2.50	2.50	**2.00**	**1.67**	**2.33**
School C (4)	**1.83**	2.50	**2.00**	**2.00**	**1.50**	**1.00**	**2.00**
School L (1)	**2.00**	ND	ND	ND	ND	ND	**2.00**
School K (2)	3.00	3.00	3.00	3.00	3.00	3.00	3.00
School I (3)	2.50	3.00	**2.33**	**2.33**	**2.33**	**2.00**	3.00
School M (2)	**1.33**	**1.00**	**2.00**	**1.00**	**1.00**	**1.00**	**2.00**
School A (2)	3.67	4.00	4.00	3.00	4.00	3.00	4.00
School H (1)	ND	ND	ND	ND	ND	ND	ND
School J (1)	3.00	ND	3.00	3.00	3.00	ND	ND
Mean	**2.31**	2.63	**2.48**	**2.35**	**2.23**	**1.81**	**2.42**
SD	0.76	0.95	0.89	0.69	1.07	0.90	0.90

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity
ND = No data
Red indicates the rounded mean rating falls below the threshold for above average fidelity.

Table B- 6. Average Effective School Practice Survey Subscale Scores for Curriculum, Assessment, and Instructional Planning, by Item, by School

		Curriculum, Assessment, and Instructional Planning (10 Items)									
School (number of raters)	Mean*	Units of instruction	Frequency of usage of multiple data points to target learning	Frequency of differentiated instruction based on pre-test data	Frequency of re-teaching based on post-test data	Frequency of development and sharing of learning materials by collaborative teams	Frequency of WVEIS data entry timeliness	Frequency of WVEIS data entry validation	Frequency of WVEIS data report generation	Frequency of WVEIS data usage for assessment of strategies	Frequency of WVEIS data usage for planning of instructional support
School N (1)	**2.40**	3.00	**2.00**	**2.00**	3.00	**2.00**	ND	ND	ND	ND	ND
School F (2)	**1.90**	3.00	3.00	**2.00**	**2.00**	**2.00**	3.00	**1.00**	**1.00**	**1.00**	**1.00**
School D (3)	2.58	3.67	2.67	**2.33**	**2.33**	**2.33**	2.50	2.50	2.50	2.50	2.50
School C (4)	2.82	3.75	3.00	3.25	2.50	**2.33**	**2.33**	**2.00**	3.00	3.00	3.00
School L (1)	2.50	3.00	ND	**2.00**	3.00	**2.00**	ND	ND	ND	ND	ND
School K (2)	2.92	3.38	**1.50**	2.63	2.75	4.33	ND	ND	ND	ND	ND
School I (3)	2.63	3.67	**2.00**	**2.00**	3.33	3.00	3.67	**1.67**	3.33	**2.00**	**1.67**
School M (2)	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**1.00**	3.00	**2.00**	**2.00**	**2.00**	**2.00**
School A (2)	3.75	3.50	3.50	3.50	3.50	3.50	4.00	4.00	4.00	4.00	4.00
School H (1)	3.00	3.00	3.00	3.00	3.00	ND	ND	ND	ND	ND	ND
School J (1)	3.20	4.00	3.00	3.00	3.00	3.00	ND	ND	ND	ND	ND
Mean	2.70	3.27	2.57	2.52	2.77	2.55	3.08	**2.19**	2.64	**2.42**	**2.36**
SD	0.53	0.55	0.64	0.58	0.50	0.94	0.65	1.01	1.06	1.02	1.06

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity
ND = No data
Red indicates the rounded mean rating falls below the threshold for above average fidelity.

Table B- 7. Average Effective School Practice Survey Subscale scores for Classroom Instruction, by Item, by School

		Classroom instruction (15 Items)														
School (Number of Raters)	Mean*	Development of weekly lesson plans	Student progress assessment	Differentiation of assignments	Instructional methods	Frequency of direct and thorough explanation	Review methods and frequency	Frequency of key concept summary	Frequency of opportunity for peer interaction	Frequency of encouragement to paraphrase, summarize, and relate	Frequency of encouragement of students to self-assess	Frequency of feedback on homework	Frequency of communication with parents	Maintenance of well-organized materials	Positive reinforcement of classroom rules	Frequency of engagement with all students
School N (1)	**2.46**	3.00	3.00	**2.00**	**2.00**	3.00	**2.00**	3.00	**2.00**	**2.00**	**2.00**	ND	ND	3.00	3.00	**2.00**
School F (2)	**2.40**	3.00	**2.00**	2.50	2.50	2.50	**2.00**	2.50	2.50	**2.00**	2.50	**2.00**	**2.00**	3.00	3.00	**2.00**
School D (3)	2.64	4.00	3.00	2.67	**2.33**	2.67	**2.33**	3.00	**2.33**	**2.33**	**2.33**	2.67	**2.33**	3.00	**2.33**	**2.33**
School C (4)	2.55	3.50	2.75	**2.00**	**2.00**	2.67	**2.00**	3.00	**2.25**	2.75	2.50	3.00	**2.00**	2.67	2.67	2.50
School L (1)	**2.38**	3.00	3.00	ND	**2.00**	3.00	**2.00**	3.00	**2.00**	**2.00**	**2.00**	**2.00**	ND	3.00	**2.00**	**2.00**
School K (2)	3.31	4.00	3.50	3.00	3.00	3.00	4.00	3.00	3.00	ND	3.00	ND	4.00	3.00	4.00	2.50
School I (3)	2.71	3.67	3.33	**2.00**	2.50	3.00	3.00	**2.00**	3.33	**2.00**	**2.00**	3.00	**2.33**	3.00	2.50	3.00
School M (2)	**1.97**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	2.50	2.50	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**1.00**	**1.50**
School A (2)	3.63	4.00	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	4.00	4.00	4.00	3.50	3.50
School H (1)	2.54	4.00	3.00	**2.00**	**2.00**	3.00	**2.00**	3.00	**2.00**	**2.00**	**2.00**	ND	ND	3.00	3.00	**2.00**
School J (1)	3.57	4.00	4.00	3.00	4.00	4.00	3.00	4.00	3.00	3.00	3.00	4.00	ND	4.00	4.00	3.00
Mean	2.74	3.47	3.01	**2.47**	2.53	2.94	2.53	2.95	2.58	**2.36**	**2.44**	2.83	2.67	3.06	2.82	**2.39**
SD	0.53	0.65	0.61	0.55	0.69	0.52	0.73	0.52	0.54	0.54	0.52	0.84	0.92	0.55	0.88	0.58

*1 = Little or no fidelity, 2 = Limited fidelity, 3 = Above average fidelity, and 4 = Strong fidelity

ND = No data

Red indicates the rounded mean rating falls below the threshold for above average fidelity.

Appendix C. Detailed RM ANOVA Model Statistics

Table C- 1. Grade 3-to-4 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
Mathematics							
TIME	49741.282	1	49741.282	61.645	.000	.091	1.00
TIME * GROUP	8622.957	1	8622.957	10.686	.001	.017	.90
Error (TIME)	499472.666	619	806.903				N/A
Reading/Language Arts							
TIME	88803.339	1	88803.339	177.053	.000	.222	1.00
TIME * GROUP	2680.672	1	2680.672	5.345	.021	.009	.64
Error (TIME)	310468.252	619	501.564				N/A

Table C- 2. Grade 3-to-4 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
Mathematics							
GROUP	2617.775	1	2617.775	.791	.374	.001	.144
Error	2047864.630	619	3308.344				N/A
Reading/Language Arts							
GROUP	1738.575	1	1738.575	.556	.456	.001	.115
Error	1936522.526	619	3128.469				N/A

Table C- 3. Grade 4-to-5 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	119224.315	1	119224.315	130.200	.000	.209	1.00
TIME * GROUP	26.161	1	26.161	.029	.866	.000	.053
Error (TIME)	450524.313	492	915.700				N/A
			Reading/Language Arts				
TIME	58307.623	1	58307.623	112.868	.000	.187	1.00
TIME * GROUP	131.347	1	131.347	.254	.614	.001	.079
Error (TIME)	254166.232	492	516.598				N/A

Table C- 4. Grade 4-to-5 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	4638.134	1	4638.134	1.026	.312	.002	.173
Error	2224745.299	492	4521.840				N/A
			Reading/Language Arts				
GROUP	2159.312	1	2159.321	.597	.440	.001	.120
Error	1779835.152	492	3617.551				N/A

Table C- 5. 5-to-6 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	19082.800	1	19082.800	24.902	.000	.062	.999
TIME * GROUP	1392.141	1	1392.141	1.817	.179	.005	.270
Error (TIME)	288895.783	377	766.302				N/A
			Reading/Language Arts				
TIME	20508.949	1	20508.949	36.232	.000	.087	1.00
TIME * GROUP	2977.354	1	2977.354	5.260	.022	.014	.628
Error (TIME)	213963.362	378	566.041				N/A

Table C- 6. Grade 5-to-6 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	2780.219	1	2780.219	.640	.424	.002	.126
Error	1636910.084	377	4341.937				N/A
			Reading/Language Arts				
GROUP	2602.286	1	2602.286	.724	.395	.002	.136
Error	1359177.377	378	3595.707				N/A

Table C- 7. Grade 6-to-7 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	188953.128	1	188953.128	256.059	.000	.148	1.00
TIME * GROUP	133.677	1	133.677	.181	.670	.000	.071
Error (TIME)	1087705.274	1474	737.928				N/A
			Reading/Language Arts				
TIME	268499.064	1	268499.064	506.845	.000	.256	1.00
TIME * GROUP	2785.190	1	2785.190	5.258	.022	.004	.630
Error (TIME)	781374.544	1475	529.745				N/A

Table C- 8. Grade 6-to-7 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	1776.957	1	1776.957	.492	.483	.000	.108
Error	5321350.531	1474	3610.143				N/A
			Reading/Language Arts				
GROUP	24215.121	1	24215.121	7.316	.007	.005	.771
Error	4882190.232	1475	3309.959				N/A

Table C- 9. Grade 7-to-8 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	1022.102	1	1022.102	1.042	.308	.001	.175
TIME * GROUP	188.141	1	188.141	.192	.662	.000	.072
Error (TIME)	1399091.830	1426	981.130				N/A
			Reading/Language Arts				
TIME	93382.345	1	93382.346	192.250	.000	.119	1.00
TIME * GROUP	455.790	1	455.790	.938	.333	.001	.162
Error (TIME)	693141.839	1427	485.734				N/A

Table C- 10. Grade 7-to-8 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	7864.292	1	7864.292	1.643	.200	.001	.249
Error	6826589.791	1426	4787.230				N/A
			Reading/Language Arts				
GROUP	22135.893	1	22135.893	7.055	.008	.005	.756
Error	4477331.951	1427	3137.584				N/A

Table C- 11. Grade 8-to-9 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	20605.306	1	20605.306	15.141	.000	.086	.972
TIME * GROUP	2477.577	1	2477.577	1.821	.179	.011	.269
Error (TIME)	217744.673	160	1360.904				N/A
			Reading/Language Arts				
TIME	5353.859	1	5353.859	6.792	.010	.042	.736
TIME * GROUP	49.555	1	49.555	.063	.802	.000	.057
Error (TIME)	122972.081	156	788.283				N/A

Table C- 12. Grade 8-to-9 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	56096.549	1	1	12.030	.001	.070	.932
Error	746077.923	160	160				N/A
			Reading/Language Arts				
GROUP	29702.500	1	29702.500	7.283	.008	.045	.765
Error	636229.009	156	4078.391				N/A

Table C- 13. Grade 9-to-10 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	20027.990	1	20027.990	12.629	.000	.019	.944
TIME * GROUP	29.734	1	29.734	.019	.891	.000	.052
Error (TIME)	1062504.009	670	1585.827				N/A
			Reading/Language Arts				
TIME	6287.822	1	6287.822	5.214	.023	.008	.626
TIME * GROUP	11153.216	1	11153.216	9.248	.002	.014	.859
Error (TIME)	810429.871	672	1205.997				N/A

Table C- 14. Grade 9-to-10 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	7179.155	1	7179.155	1.741	.187	.003	.261
Error	2763090.184	670	4124.015				N/A
			Reading/Language Arts				
GROUP	10632.654	1	10632.654	1.994	.158	.003	.292
Error	3584112.035	672	5333.500				N/A

Table C- 15. Grade 10-to-11 Cohort Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
TIME	53822.980	1	53822.980	48.021	.000	.077	1.00
TIME * GROUP	98.613	1	98.613	.088	.767	.000	.060
Error (TIME)	644475.360	575	1120.827				N/A
			Reading/Language Arts				
TIME	10061.833	1	10061.833	8.143	.004	.014	.813
TIME * GROUP	3400.040	1	3400.040	2.752	.098	.005	.381
Error (TIME)	714223.839	578	1235.681				N/A

Table C- 16. Grade 10-to-11 Cohort Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2	Observed Power
			Mathematics				
GROUP	630.648	1	630.648	.185	.667	.000	.071
Error	1958069.914	575	3405.339				N/A
			Reading/Language Arts				
GROUP	10519.314	1	10519.314	1.455	.228	.003	.226
Error	4178231.406	578	7228.774				N/A

Appendix D. Subgroup Trends in SIG and non-SIG Comparison Schools

Table D- 1 provides an overview of for which schools there were enough students continuously enrolled in order for us to analyze student achievement data by subgroup. Figure D- 1 through Figure D- 25 in this appendix present the trends in SY2009 to SY2011 achievement for each subgroup. For occasions where subgroup data were available for a non-SIG comparison school, but not the matched SIG schools, we do not present the data.

Table D- 1. Subgroup Analysis by School

School	FRPL eligibility	Special education eligibility	Black or African American[2]
School U	YES	YES	YES
School X	YES[1]		
School Q	YES		
School B	YES		
School N	YES	YES	
School W	YES	YES	
School E	YES		
School F	YES		
School D	YES		
School AB	YES		
School C	YES	YES	
School AC	YES		
School L	YES	YES	
School V	YES	YES	
School Y	YES		
School P	YES	YES	YES
School K	YES		
School I	YES	YES	YES
School M	YES	YES	YES
School Z	YES		
School R	YES		
School O	YES	YES	YES
School A	YES		
School T	YES	YES	
School H	YES	YES	YES
School AA	YES		
School J	YES	YES	
School S	YES		

[1] = This school was not tested during SY2009.

[2] = The Black or African American subgroup was the only racial/ethnic subgroup for which more than 50 students were available for analysis.

Achievement Trends for Free or Reduced Price Lunch Eligible Students

All Cohort I SIG schools had enough students continuously enrolled from SY2009 to SY2011 who were eligible for free and reduced price lunch for us to examine achievement trends for the subgroup. Figure D- 1 through Figure D- 14 present trends for these schools.

Figure D- 1. Free and Reduced Price Lunch Eligible Proficiency Trend for School I and School P

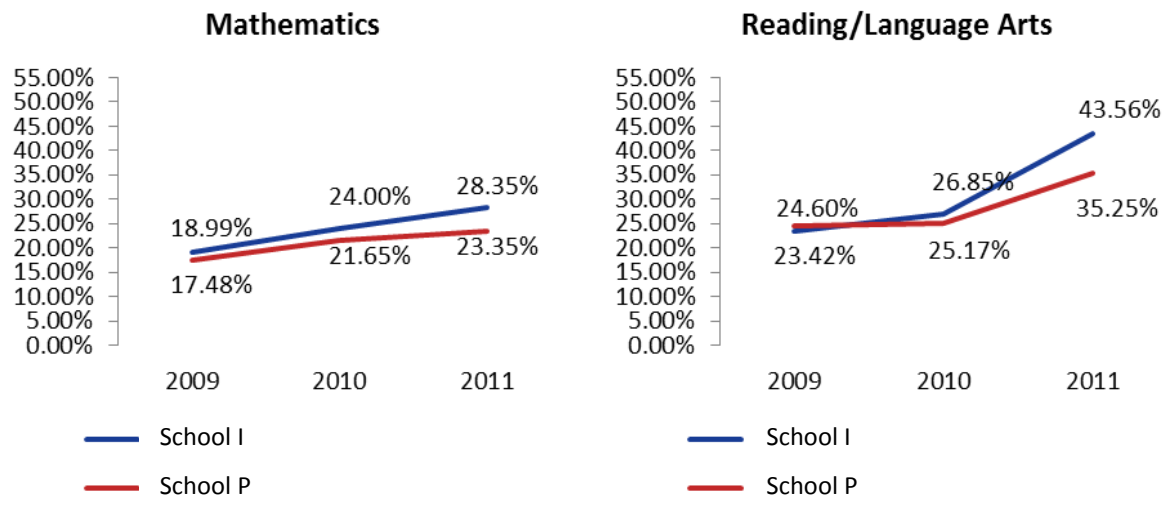


Figure D- 2. Free and Reduced Price Lunch Eligible Proficiency Trend for N and School Q

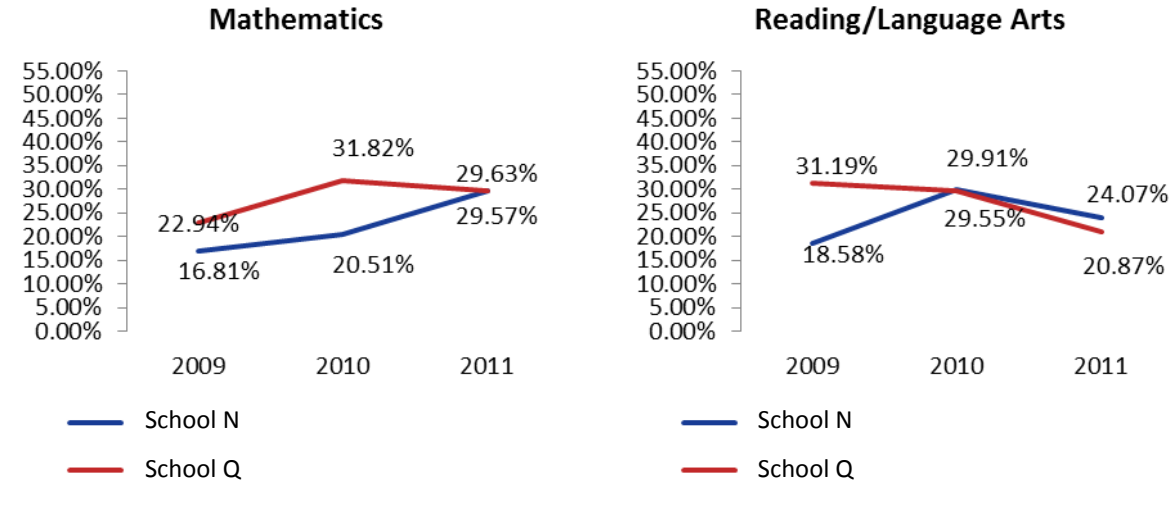


Figure D- 3. Free and Reduced Price Lunch Eligible Proficiency Trend for School K and School R

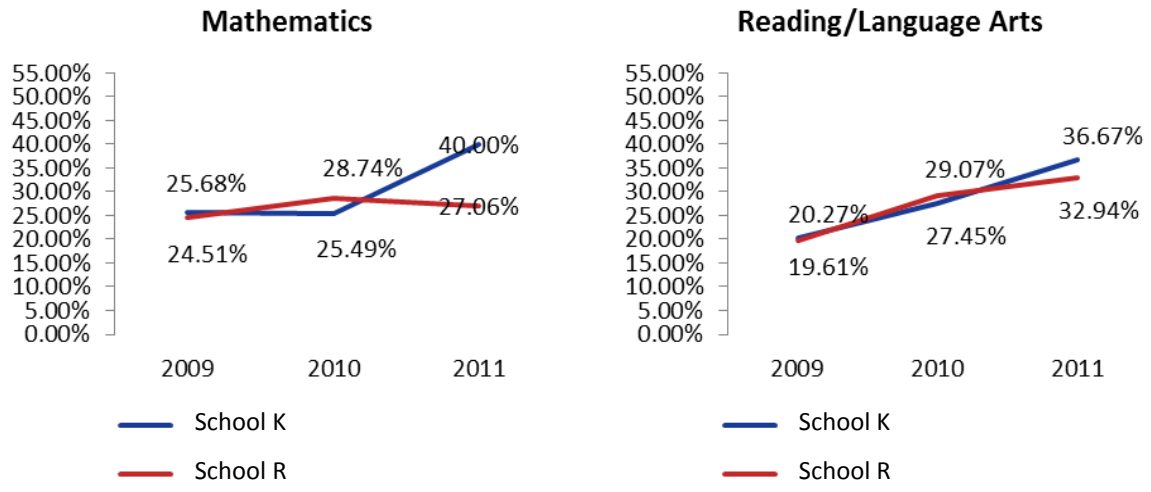


Figure D- 4. Free and Reduced Price Lunch Eligible Proficiency Trend for School B and School S

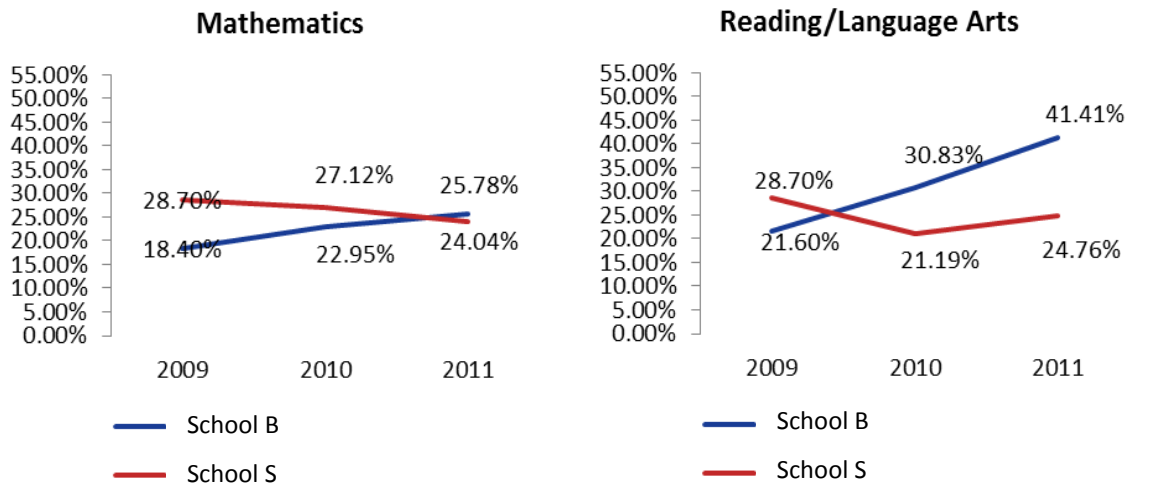


Figure D- 5. Free and Reduced Price Lunch Eligible Proficiency Trend for School E and School T

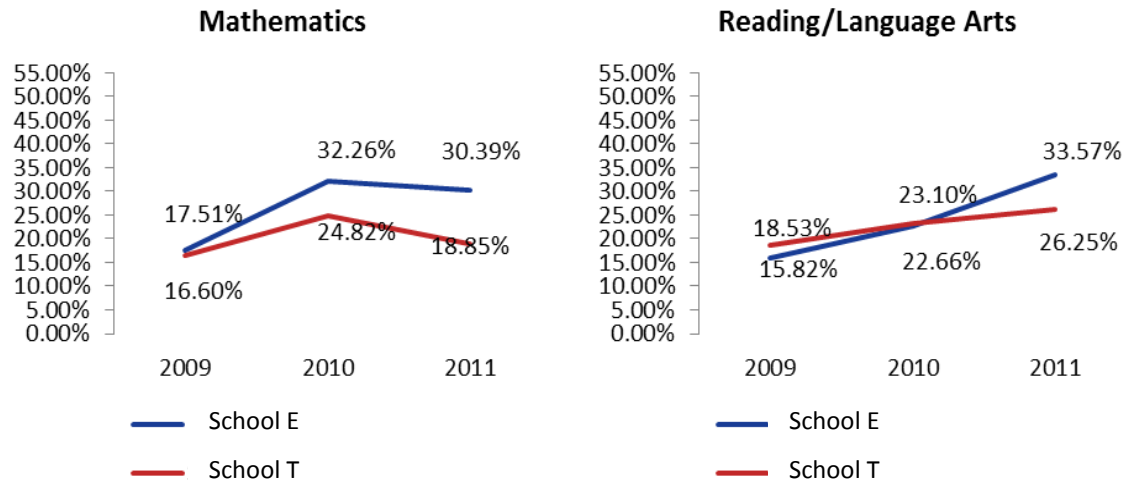


Figure D- 6. Free and Reduced Price Lunch Eligible Proficiency Trend for School H and School U

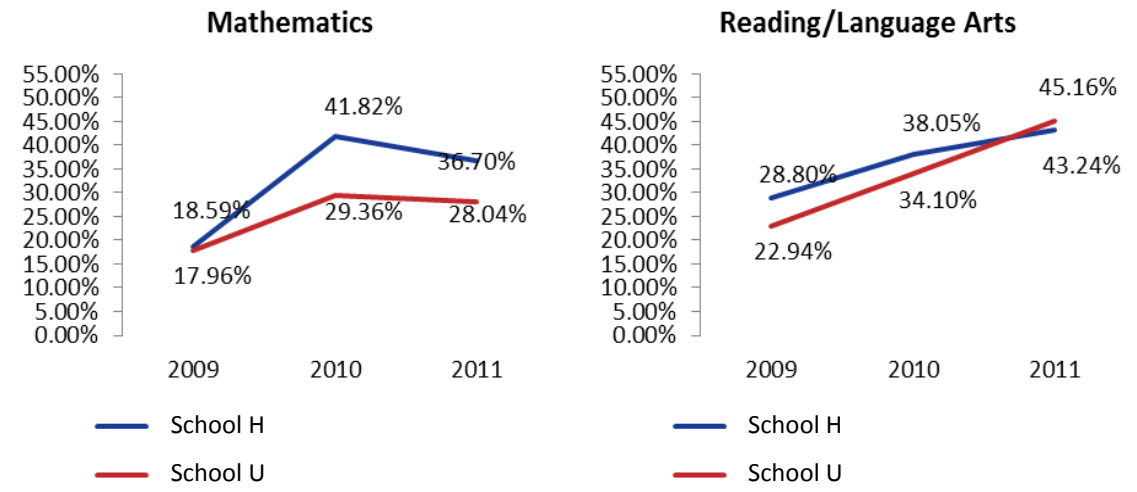


Figure D- 7. Free and Reduced Price Lunch Eligible Proficiency Trend for School O and School V

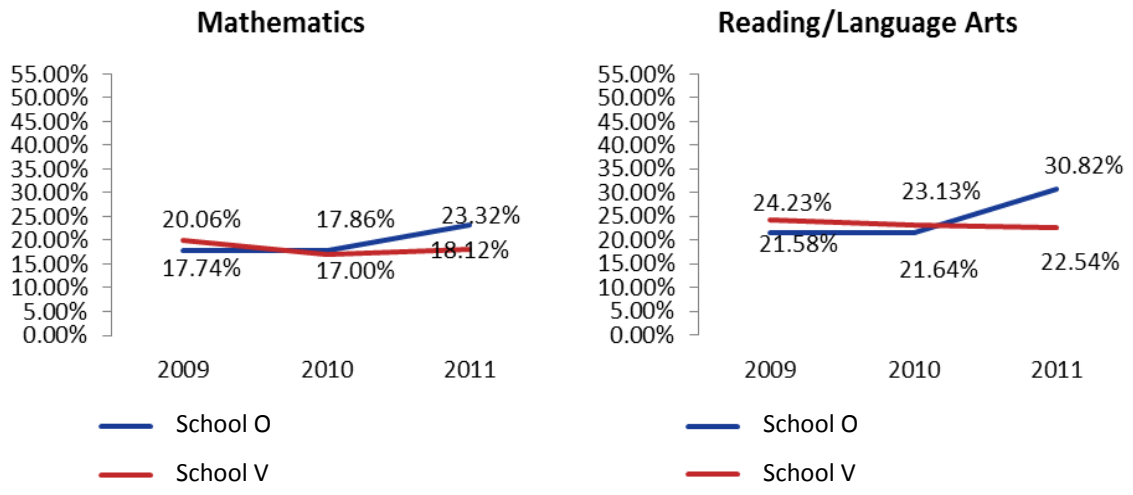


Figure D- 8. Free and Reduced Price Lunch Eligible Proficiency Trend for School L and W

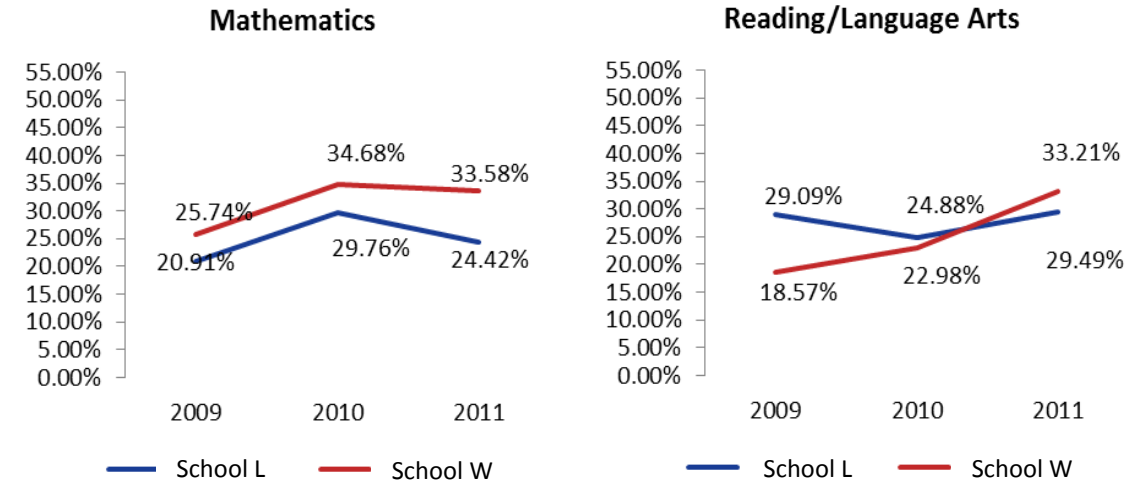


Figure D- 9. Free and Reduced Price Lunch Eligible Proficiency Trend for School J and School X

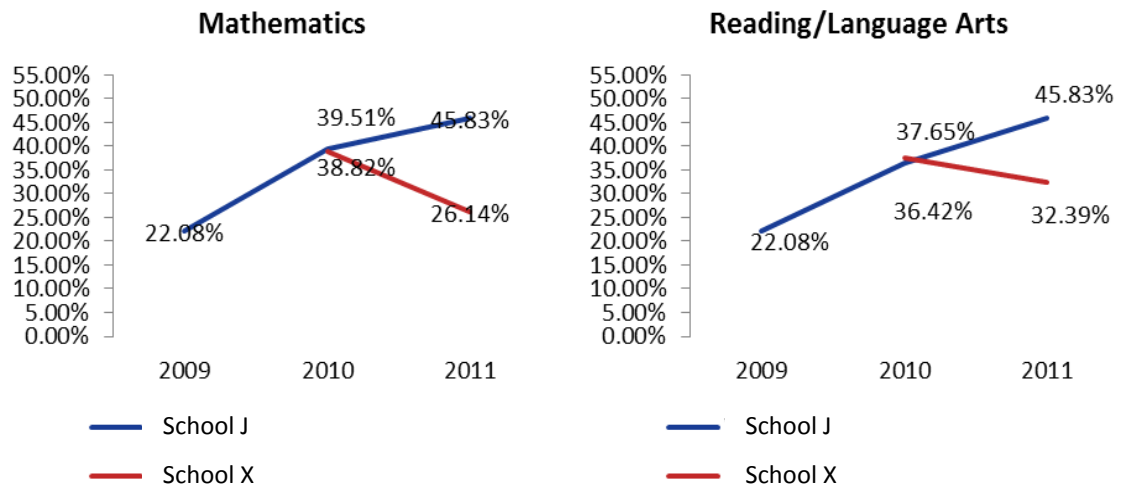


Figure D- 10. Free and Reduced Price Lunch Eligible Proficiency Trend for School C and School Y

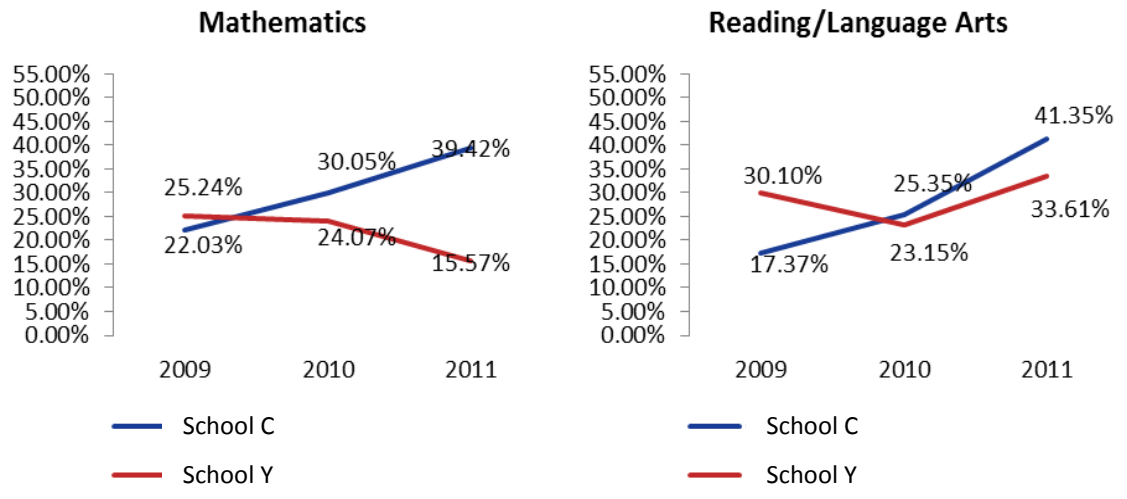


Figure D- 11. Free and Reduced Price Lunch Eligible Proficiency Trend for School A and School Z

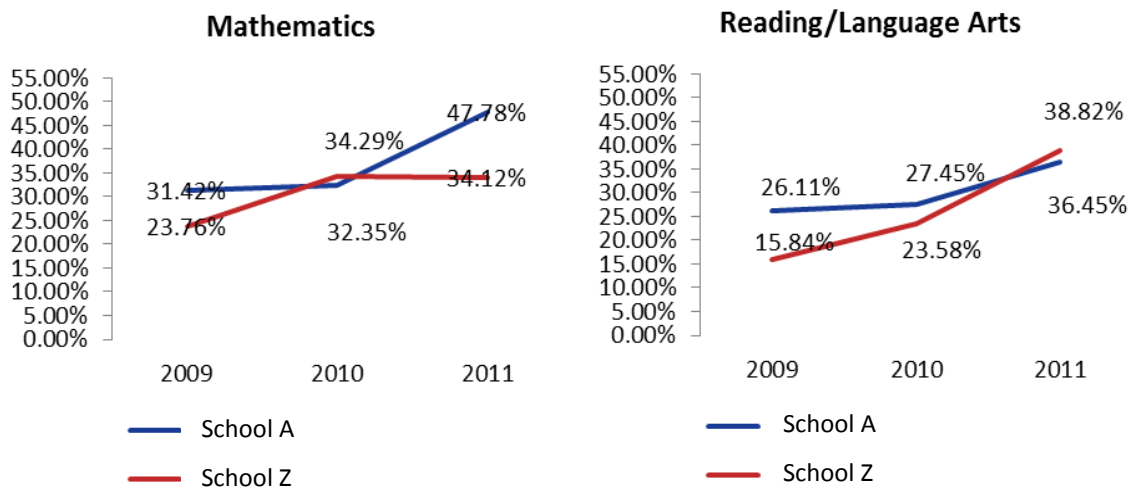


Figure D- 12. Free and Reduced Price Lunch Eligible Proficiency Trend for School M and School AA

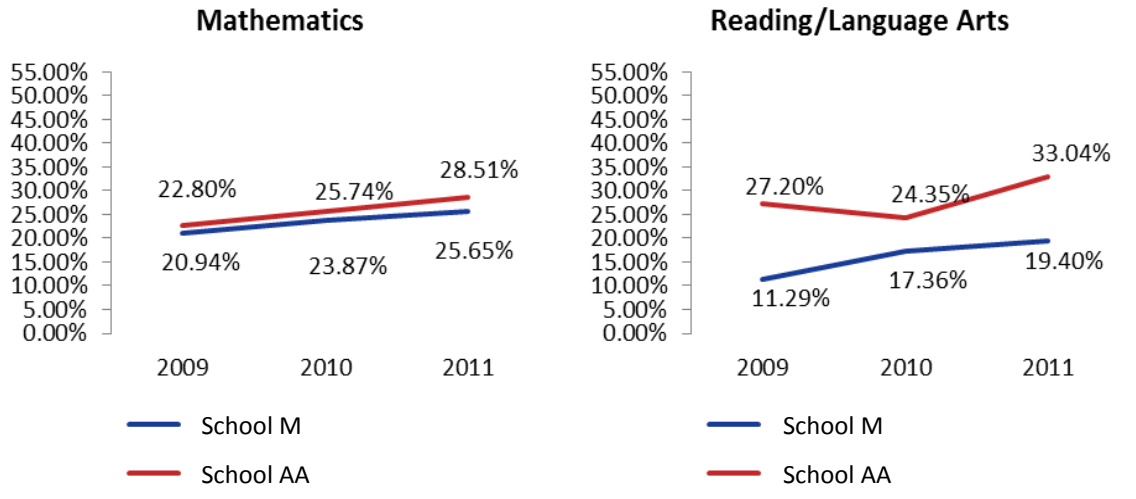


Figure D- 13. Free and Reduced Price Lunch Eligible Proficiency Trend for School D and School AB

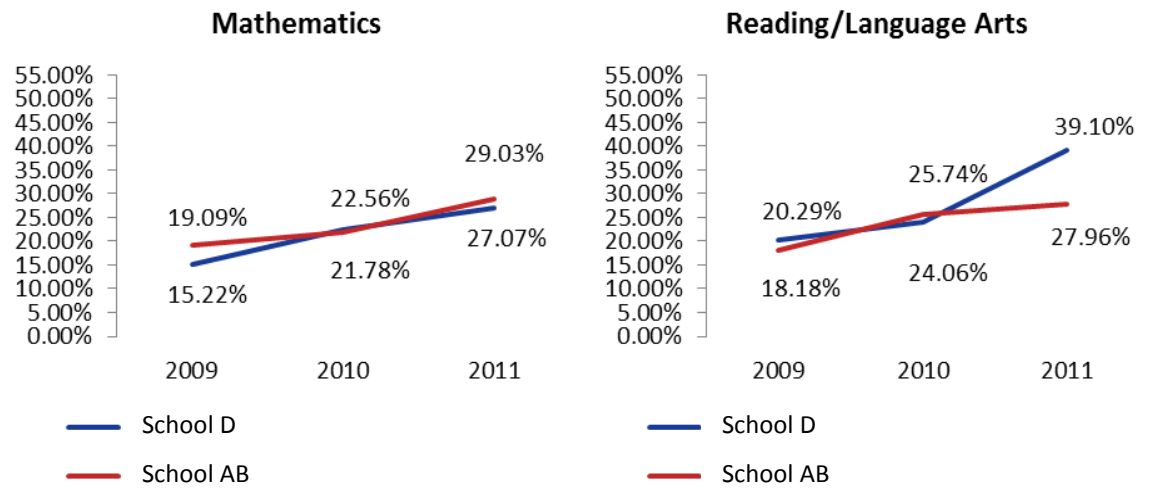
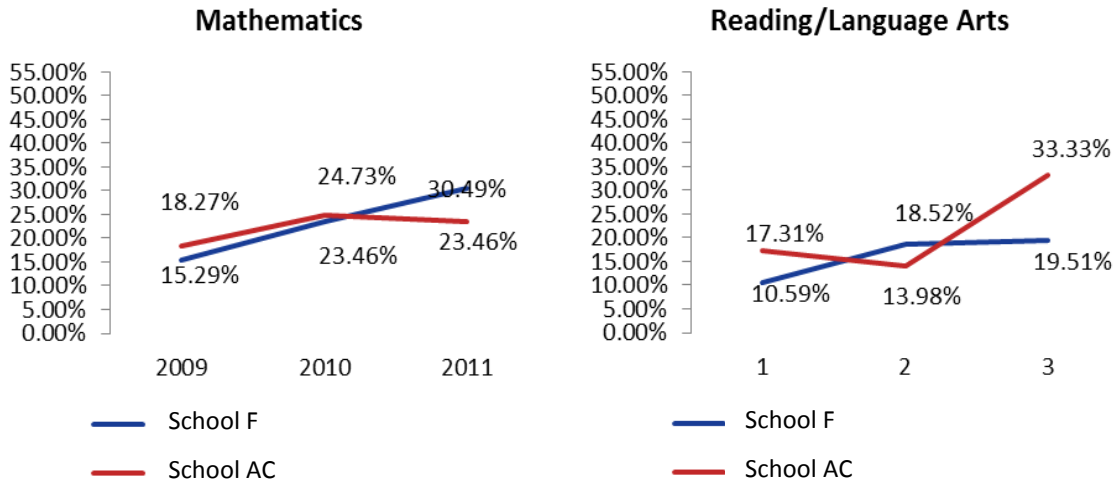


Figure D- 14. Free and Reduced Price Lunch Eligible Proficiency Trend for School F and School AC



Achievement Trends for Special Education Eligible Students

Seven Cohort I SIG schools had enough Special Education Eligible Students continuously enrolled from SY2009 to SY2011 for us to examine achievement trends for the subgroup. However, data from matched non-SIG comparison schools were only available for four of these schools. Figure D- 15 through Figure D- 21 present trends for these schools.

Figure D- 15. Special Education Eligible Proficiency Trend for School I and School P

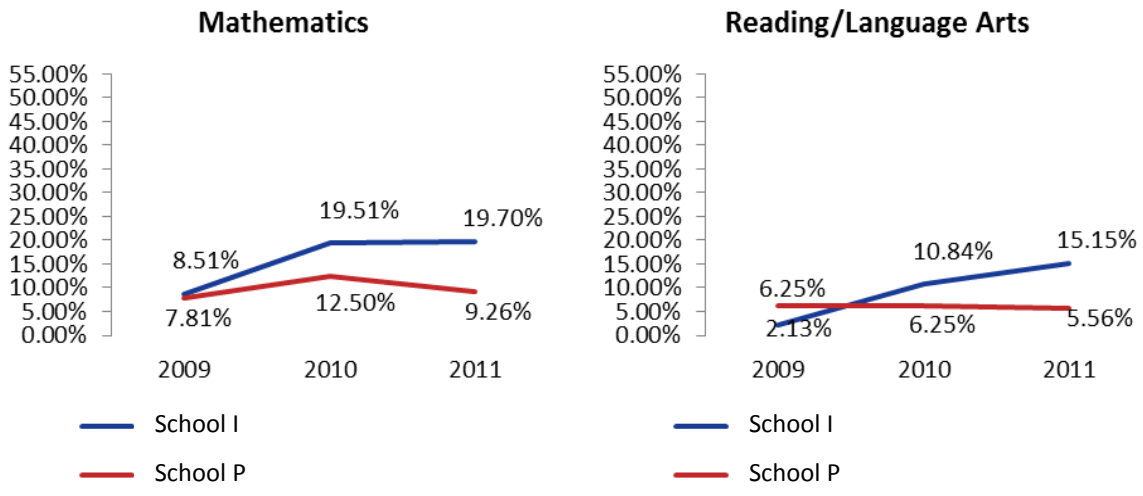


Figure D- 16. Special Education Eligible Proficiency Trend for School H and U

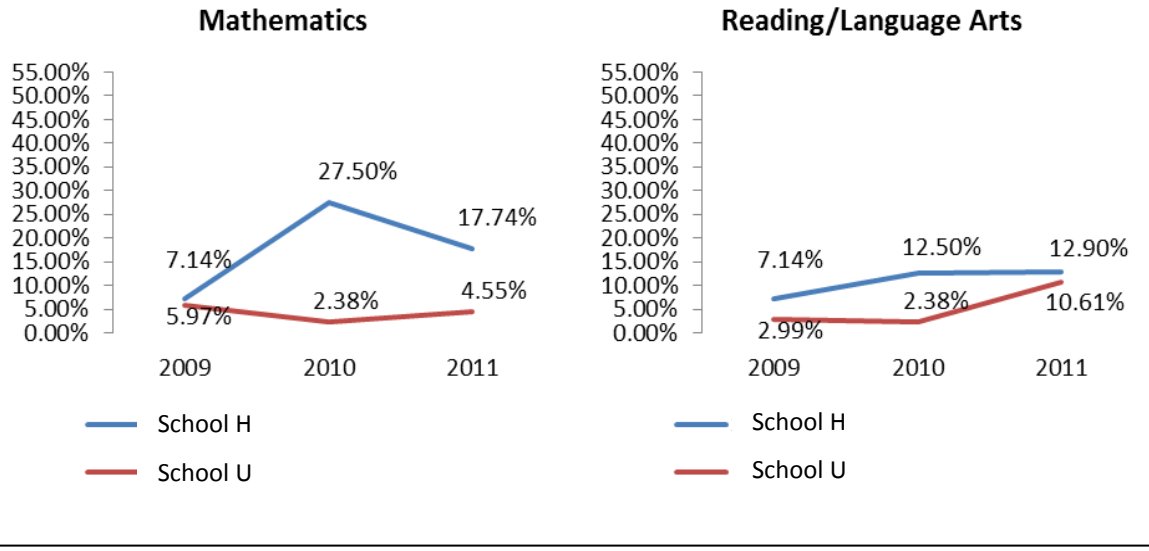


Figure D- 17. Special Education Eligible Proficiency Trend for School O and School V

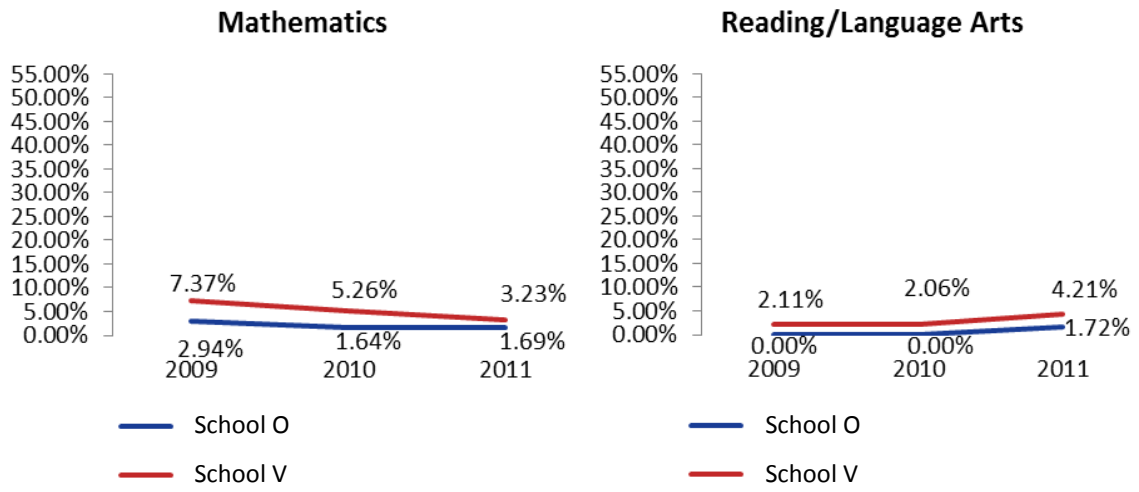


Figure D- 18. Special Education Eligible Proficiency Trend for School L and W

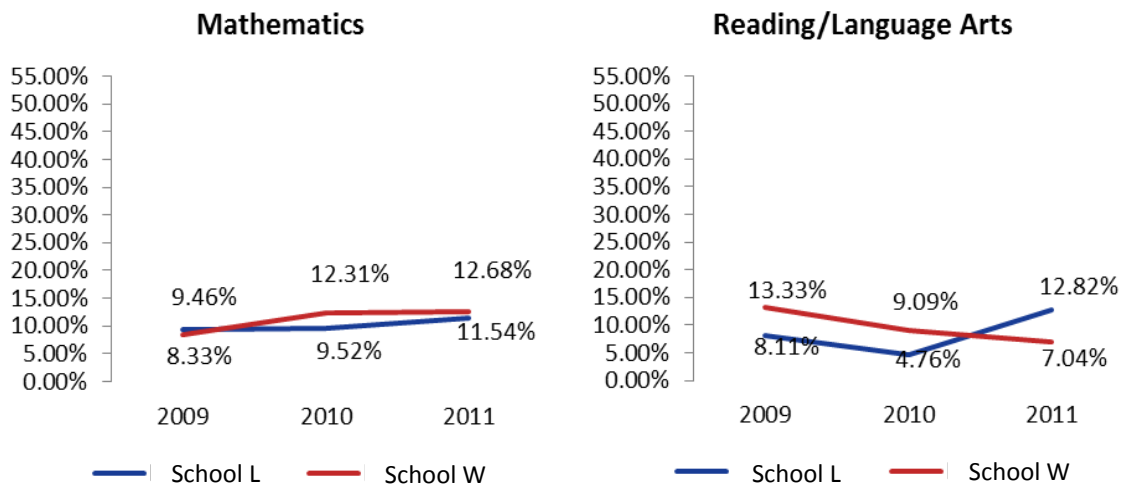


Figure D- 19. Special Education Eligible Proficiency Trend for School J

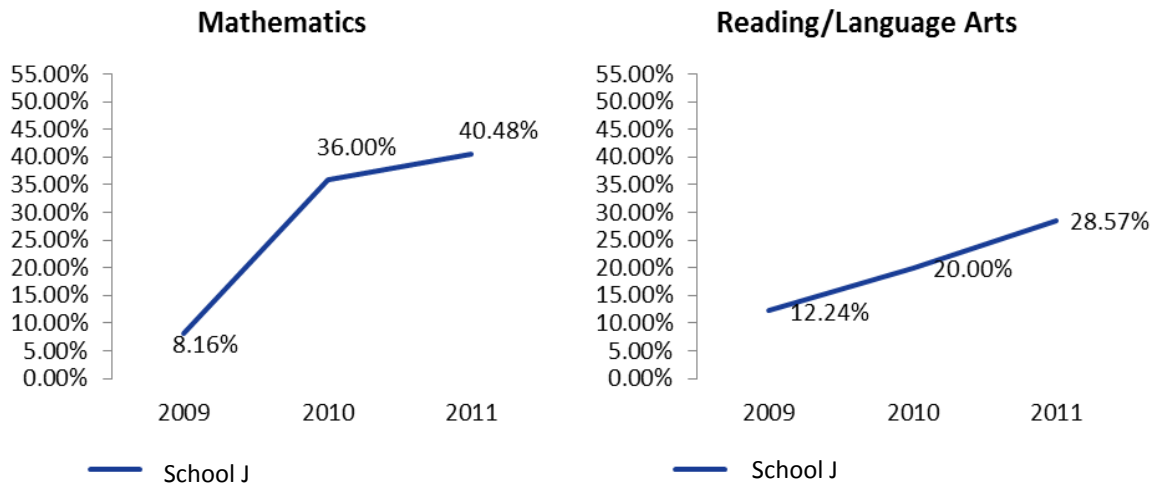


Figure D- 20. Special Education Eligible Proficiency Trend for School C

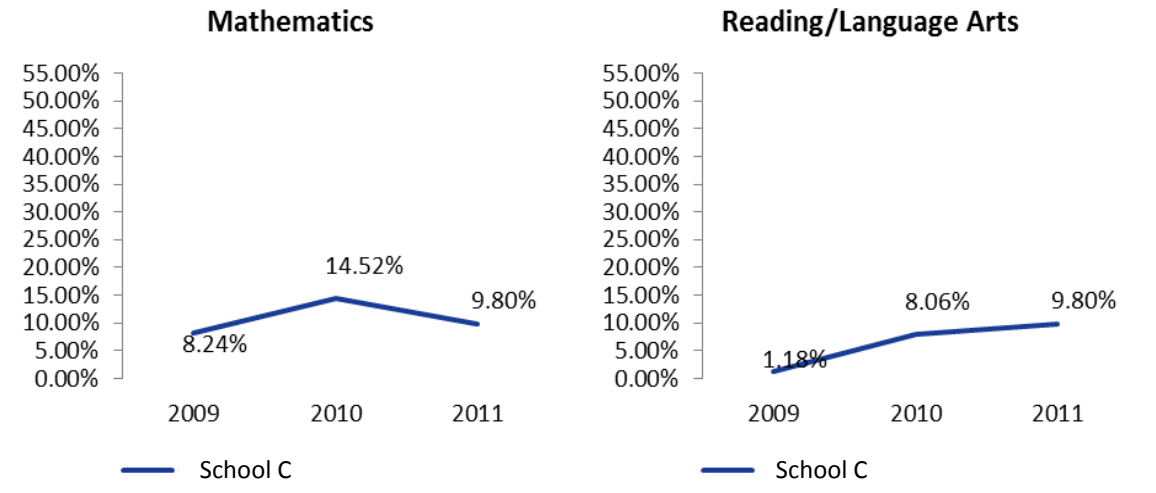
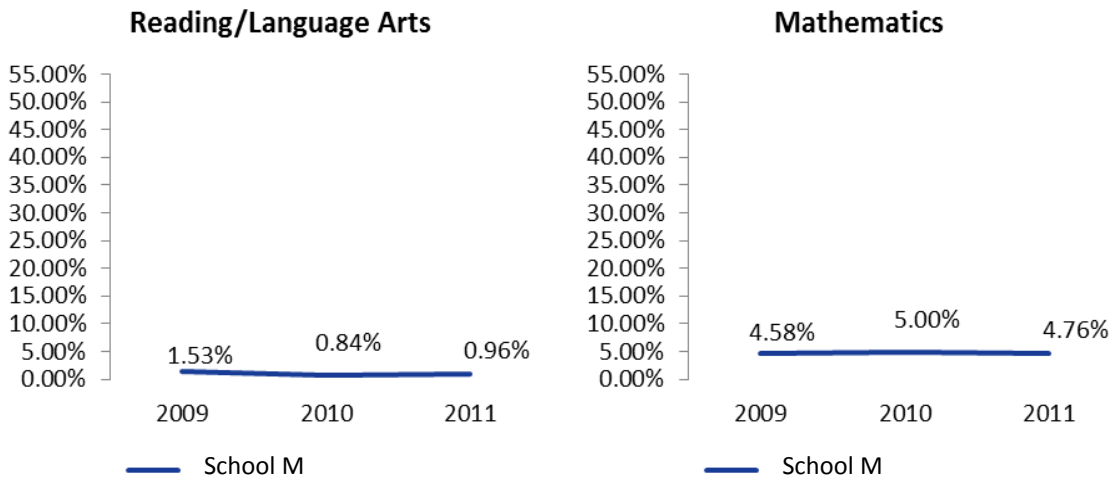


Figure D- 21. Special Education Eligible Proficiency Trend for School M



Achievement Trends for Black or African American Students

Four Cohort I SIG schools had enough Black or African American Students continuously enrolled from SY2009 to SY2011 for us to examine achievement trends for the subgroup. However, data from matched non-SIG comparison schools were only available for two of these schools. Figure D- 22 through Figure D- 25 present trends for these schools.

Figure D- 22. Black or African American Subgroup Proficiency Trend for School I and School P

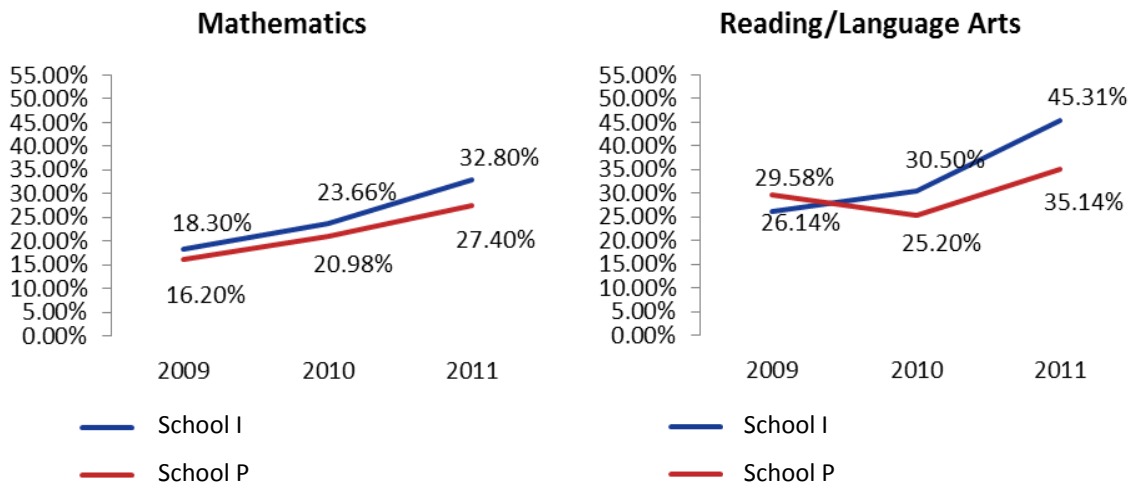


Figure D- 23. Black or African American Subgroup Proficiency Trend for School H and School U

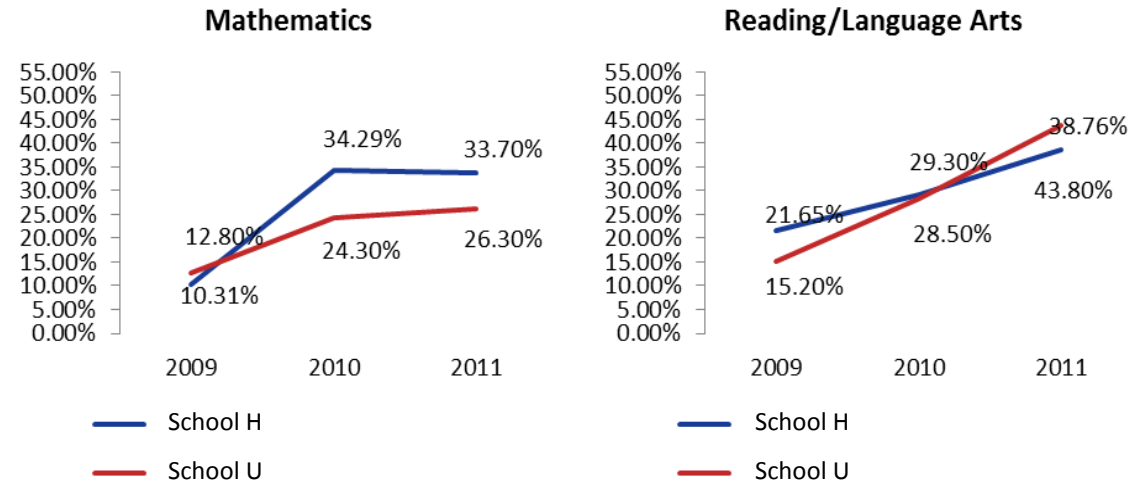


Figure D- 24. Black or African American Subgroup Proficiency Trend for School O

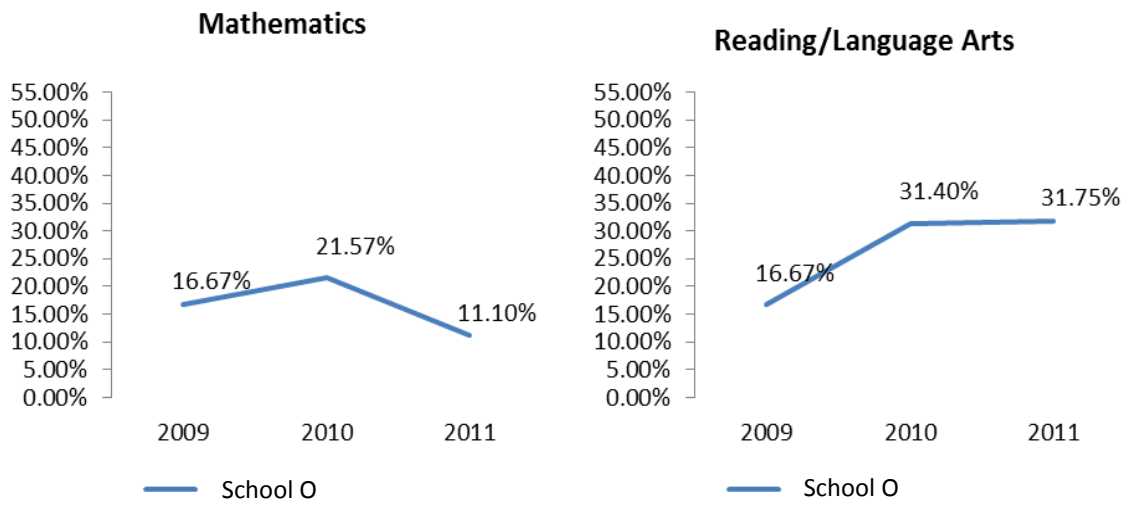
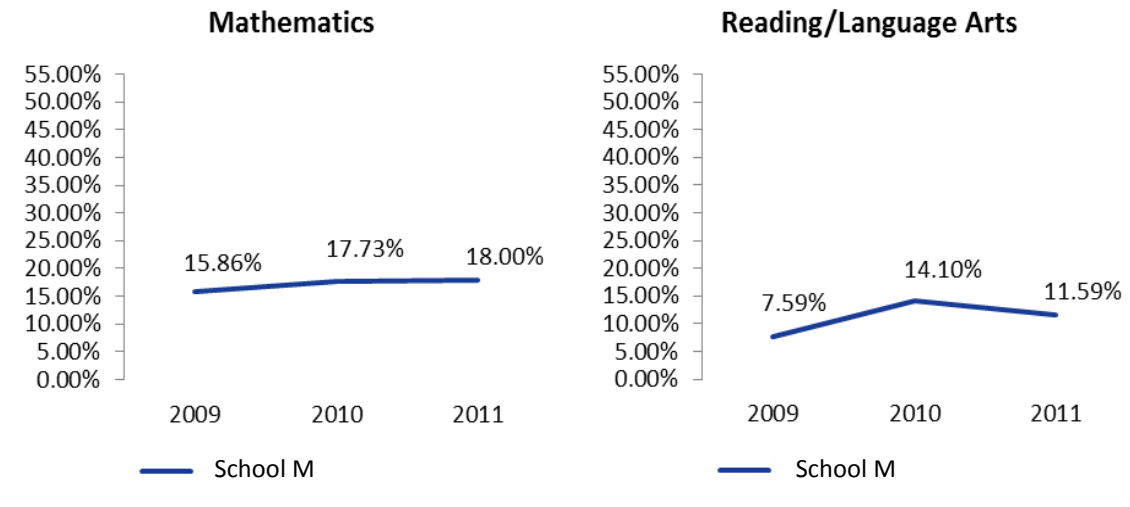


Figure D- 25. Black or African American Subgroup Proficiency Trend for School M



Appendix E. School Culture Survey Tables

Tables begin on next page.

Table E- 1. School Culture Survey Results: Average Subscale Scores for Collaborative Leadership by Item by School

		Collaborative Leadership (11 Items)										
	School Mean**	Leaders value teachers' ideas	Leaders in this school trust the professional judgments of teachers.	Leaders take time to praise teachers that perform well.	Teachers are involved in the decision-making process	Leaders in our school facilitate teachers working together.	Teachers are kept informed on current issues in the school.	My involvement in policy or decision making is taken seriously	Teachers are rewarded for experimenting with new ideas and techniques.	Leaders support risk-taking and innovation in teaching.	Administrators protect instruction and planning time.	Teachers are encouraged to share ideas.
School B	3.97	4.33	4.24	3.83	3.88	3.83	3.83	3.56	3.50	3.94	4.44	4.22
School N	3.71	4.00	3.67	3.47	3.47	3.93	3.73	3.60	3.67	3.47	3.73	4.13
School E	3.63	3.61	3.54	3.86	3.29	3.96	3.82	3.43	3.25	3.68	3.43	4.14
School F	**2.77**	**2.56**	**2.74**	**2.67**	**2.41**	**2.89**	**2.81**	**2.62**	**2.60**	**2.76**	3.04	3.32
School D	3.61	3.79	3.76	3.55	3.38	3.76	3.52	3.59	3.11	3.52	3.69	4.00
School C*	3.98	3.79	4.17	4.04	3.83	4.29	3.79	3.83	3.46	4.04	4.38	4.21
School L	3.20	3.12	2.83	**2.96**	2.96	3.56	3.83	3.08	**2.79**	**2.74**	3.58	3.75
School K	3.71	4.00	3.67	3.47	3.47	3.93	3.73	3.60	3.67	3.47	3.73	4.13
School I	3.11	3.28	3.30	**2.78**	3.08	3.36	3.11	**2.89**	**2.76**	3.24	**2.73**	3.68
School M	**2.83**	**2.81**	**2.72**	**2.98**	**2.37**	3.42	**2.58**	**2.35**	**2.60**	3.12	**2.67**	3.47
School O	3.76	3.98	3.73	3.81	3.33	3.57	3.81	3.48	3.50	4.02	3.81	4.33
School A	4.14	4.24	4.24	4.19	3.95	4.38	3.71	3.86	4.05	4.14	4.33	4.48
School H	3.95	4.26	4.19	3.80	3.83	4.19	3.90	3.79	3.31	4.00	3.88	4.26
School J*	4.02	4.07	4.14	3.86	3.93	4.45	4.03	3.79	3.41	3.93	4.28	4.28
Mean	3.60	3.70	3.64	3.52	3.37	3.82	3.59	3.39	3.26	3.58	3.69	4.03
(SD)	(0.45)	(0.56)	(0.55)	(0.49)	(0.52)	(0.44)	(0.44)	(0.47)	(0.44)	(0.47)	(0.58)	(0.34)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Scores in **red** indicate school respondents' mean disagreement with statement indicating presence of collaborative leadership.

Table E- 2. School Culture Survey Results: Average Subscale Scores for Teacher Collaboration by Item by School

		Teacher Collaboration (6 Items)					
School	Mean**	Teachers have opportunities for dialogue and planning across grades and subjects.	Teachers spend considerable time planning together.	Teachers take time to observe each other teaching.	Teachers are generally aware of what other teachers are teaching.	Teachers work together to develop and evaluate programs and projects.	Teaching practice disagreements are voiced openly and discussed.
School B	3.45	3.72	3.44	**2.88**	3.78	3.61	3.28
School N	3.66	3.87	3.67	**2.86**	3.93	4.07	3.47
School E	3.32	3.57	3.21	**2.57**	3.86	3.46	3.30
School F	**2.78**	3.00	3.12	**2.35**	**2.88**	**2.77**	**2.54**
School D	3.01	3.52	**2.90**	**2.36**	3.14	3.31	**2.82**
School C*	3.55	3.91	3.88	**2.63**	3.46	4.00	3.46
School L	3.13	3.28	3.38	**2.76**	3.60	3.50	**2.24**
School K	3.66	3.87	3.67	**2.86**	3.93	4.07	3.47
School I	**2.90**	3.68	**2.73**	**2.19**	**2.78**	3.22	**2.81**
School M	**2.71**	**2.77**	**2.49**	**2.33**	**2.84**	**2.93**	**2.93**
School O	3.26	3.83	**2.76**	**2.67**	3.33	3.67	3.27
School A	4.27	4.00	3.81	4.10	3.76	6.29	3.67
School H	3.38	4.14	3.40	**2.43**	3.36	3.62	3.36
School J*	3.79	3.93	4.52	**2.69**	3.97	4.10	3.52
Mean	3.35	3.65	3.35	**2.69**	3.47	3.76	3.15
(SD)	(0.43)	(0.39)	(0.54)	(0.46)	(0.43)	(0.84)	(0.42)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Scores in **red** indicate school respondents' mean disagreement with statement indicating presence of teacher collaboration.

Table E- 3. School Culture Survey Results: Average Subscale Scores for Professional Development by Item by School

School	Mean **	Professional Development (5 Items)				
		Teachers utilize Professional networks to obtain information and resources for classroom instruction.	Teachers regularly seek ideas from seminars, colleagues, and conferences.	Professional development is valued by the faculty.	Teachers maintain a current knowledge base about the learning process.	The faculty values school improvement.
School B	3.85	4.12	3.65	3.17	4.06	4.24
School N	4.26	4.40	4.07	3.87	4.47	4.47
School E	3.89	4.18	3.89	3.39	3.93	4.07
School F	3.64	3.88	3.50	3.12	3.80	3.91
School D	3.63	3.69	3.48	3.21	3.59	4.17
School C*	4.13	4.25	4.00	3.92	4.13	4.33
School L	3.91	4.00	3.60	3.71	4.08	4.16
School K	4.26	4.40	4.07	3.87	4.47	4.47
School I	3.56	3.78	3.31	3.11	3.57	4.05
School M	3.86	3.93	3.71	3.61	3.93	4.14
School O	4.11	4.36	3.93	3.76	4.26	4.26
School A	4.39	4.38	4.52	4.33	4.24	4.48
School H	4.01	4.24	3.81	3.64	4.05	4.31
School J*	4.12	4.21	3.93	4.03	4.10	4.34
Mean	3.95	4.12	3.79	3.58	4.03	4.23
(SD)	(0.28)	(0.25)	(0.34)	(0.41)	(0.29)	(0.19)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Table E- 4. School Culture Survey Results: Average Subscale Scores for Unity of Purpose by Item by School

			Unity of Purpose (5 items)				
School		Mean**	Teachers support the mission of the school.	The school mission provides a clear sense of direction for teachers.	Teachers understand the mission of the school.	The school mission statement reflects the values of the community.	Teaching performance reflects the mission of the school.
School B		3.88	3.94	4.06	4.06	3.33	4.00
School N		4.12	4.20	4.13	4.20	3.73	4.27
School E		3.85	3.89	4.11	4.21	3.18	3.86
School F		3.29	3.41	3.30	3.35	3.12	3.28
School D		3.86	4.03	3.89	4.07	3.43	3.86
School C*		4.00	4.33	4.08	4.04	3.67	3.88
School L		3.96	4.24	3.96	4.00	3.72	3.88
School K		4.12	4.20	4.13	4.20	3.73	4.27
School I		3.59	3.84	3.46	3.81	3.41	3.43
School M		3.47	3.81	3.37	3.53	3.12	3.51
School O		3.87	4.10	3.93	3.98	3.36	4.00
School A		4.33	4.52	4.38	4.52	3.81	4.43
School H		4.02	4.17	4.17	4.00	3.83	3.90
School J*		4.21	4.24	4.38	4.31	3.97	4.17
	Mean	3.90	4.07	3.95	4.02	3.53	3.91
	(SD)	(0.29)	(0.28)	(0.34)	(0.30)	(0.28)	(0.33)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Table E- 5. School Culture Survey Results: Average Subscale Scores for Collegial Support by Item by School

		Collegial Support (4 items)			
School	Mean**	Teachers trust each other.	Teachers are willing to help out whenever there is a problem.	Teachers' ideas are valued by other teachers.	Teachers work cooperatively in groups.
School B	3.86	3.56	4.50	3.67	3.72
School N	4.12	3.73	4.27	4.13	4.33
School E	3.88	3.71	4.11	3.93	3.81
School F	3.61	3.11	3.84	3.81	3.68
School D	3.74	3.38	4.28	3.72	3.57
School C*	3.99	3.79	4.17	4.04	3.96
School L	3.93	3.72	4.00	4.04	3.96
School K	4.12	3.73	4.27	4.13	4.33
School I	3.55	3.14	3.89	3.62	3.54
School M	3.43	3.07	3.69	3.37	3.60
School O	4.07	3.79	4.40	4.12	3.95
School A	4.50	4.43	4.71	4.52	4.33
School H	4.02	3.86	4.21	3.98	4.05
School J*	3.95	3.45	4.28	3.79	4.28
Mean	3.91	3.60	4.19	3.92	3.94
(SD)	(0.27)	(0.36)	(0.27)	(0.28)	(0.29)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Table E- 6. School Culture Survey Results: Average Subscale Scores for Learning Partnership by Item by School

School	Mean**	Learning Partnership			
		Teachers and parents have common expectations for student performance.	Parents trust teachers' professional judgments.	Teachers and parents communicate frequently about student performance.	Students generally accept responsibility for their schooling.
School B	**2.89**	**2.67**	3.11	3.44	**2.33**
School N	3.58	3.33	3.47	4.07	3.47
School E	**2.86**	**2.71**	3.00	3.75	**2.00**
School F	**2.90**	**2.58**	3.30	3.36	**2.38**
School D	3.25	3.25	3.10	3.66	3.00
School C*	**2.81**	**2.63**	**2.88**	3.17	**2.58**
School L	3.29	3.13	3.40	3.80	**2.83**
School K	3.58	3.33	3.47	4.07	3.47
School I	**2.86**	**2.83**	3.05	3.22	**2.32**
School M	**2.38**	**2.39**	**2.67**	**2.26**	**2.21**
School O	**2.73**	**2.54**	**2.88**	3.19	**2.31**
School A	**2.85**	**2.71**	3.19	**2.95**	**2.52**
School H	**2.92**	**2.71**	3.02	3.31	**2.64**
School J*	3.62	3.48	3.41	4.03	3.55
Mean	3.04	**2.88**	3.14	3.45	**2.69**
(SD)	(0.37)	(0.35)	(0.25)	(0.50)	(0.51)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Scores in **red** indicate school respondents' mean disagreement with statements about learning partnership.

Table E- 7. School Culture Survey Results: Average Subscale Scores for Efficacy Factor by Item by School

		Efficacy Factor (10 Items)***									
School	Mean**	Success or failure in teaching due primarily to factors beyond my control.	I sometimes feel it is a waste of time to try to do my best as a teacher.	I am certain I am making a difference in the lives of students.	Attitudes and habits students bring to class greatly reduce their chances for academic success.	Many students are not capable of learning the material.	I feel responsible for students I teach but not for other students.	I am likely to be recognized publicly as a direct result of my students' academic success or failure.	Teachers are expected to help maintain discipline in the entire school.	The level of student misbehavior and/or substance use in this school interferes with my teaching.	I have been successful in providing the kind of education I wanted to give students in my classes this year.
School B	3.43	**2.89**	**2.06**	4.12	3.65	**2.28**	**1.88**	**2.94**	3.67	3.06	3.39
School N	3.95	**2.80**	**1.40**	4.40	**2.87**	**2.00**	**1.53**	3.57	4.47	**2.40**	4.07
School E	3.26	3.00	**1.67**	4.10	4.17	3.00	**2.14**	**2.77**	4.32	3.95	3.29
School F	3.58	3.19	**2.00**	4.04	3.00	**2.00**	**1.78**	3.24	4.19	**2.46**	**2.73**
School D	3.78	3.55	4.28	4.28	**2.86**	3.86	4.10	3.21	4.28	3.62	3.72
School C*	3.49	**2.75**	4.04	4.13	**2.67**	3.33	4.42	**2.83**	4.13	3.00	3.58
School L	3.84	3.79	4.21	4.00	3.29	4.04	4.28		3.96	4.28	3.74
School K	3.95	**2.80**	**1.40**	4.40	**2.87**	**2.00**	**1.53**	3.57	4.47	**2.40**	4.07
School I	3.50	3.22	4.30	3.89	**2.68**	4.00	4.35	**2.51**	3.97	**2.97**	3.11
School M	3.36	**2.95**	3.95	3.98	**2.14**	3.60	4.23	**2.67**	4.23	**2.63**	3.26
School O	3.65	3.13	4.19	4.05	**2.71**	3.67	4.19	3.44	4.48	3.00	3.69
School A	3.70	3.38	4.00	4.10	**2.81**	3.81	4.38	3.35	4.65	3.25	3.29
School H	3.63	3.32	4.52	4.24	**2.48**	3.79	4.17	**2.88**	4.52	**2.69**	3.71
School J*	3.69	3.14	4.48	4.10	**2.55**	4.00	3.93	**2.76**	4.38	3.79	3.79
Mean	3.63	3.14	3.32	4.13	**2.91**	3.24	3.35	3.04	4.27	3.11	3.53
(SD)	(0.21)	(0.30)	(1.27)	(0.15)	(0.51)	(0.82)	(1.24)	(0.35)	(0.27)	(0.60)	(0.37)

*Data from SY2009-2010. Data not available for SY2010-2011

**1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

***Includes reverse scored items

Scores in **red** indicate school respondents' mean disagreement with statements about the efficacy factor.

GLOBAL21

Students deserve it • The world demands it

West Virginia Department of EDUCATION

Jorea M. Marple, Ed.D.
State Superintendent of Schools