Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

The Effects of School Renaissance on Student Achievement in Two Mississippi School Districts:

A Longitudinal Quasi-Experimental Study

Center for Research in Educational Policy

The University of Memphis
325 Browning Hall
Memphis, Tennessee 38152
Toll Free: 1-866-670-6147

The Effects of School Renaissance on Student Achievement in Two Mississippi School Districts:

A Longitudinal Quasi-Experimental Study

July 2005

Steve Ross
John Nunnery
Ashli Avis
Tom Borek
Center for Research in Educational Policy

Executive Summary

The purpose of the present study was to compare the achievement of students in 14 Pascagoula schools and 9 Biloxi schools implementing School Renaissance (SR) to that of students in matched Control schools. The achievement measures to be examined were from the 2004 administration of the Mississippi Curriculum Test (MCT) for Reading, Language Arts, and Mathematics in grades 2-8 and the Mississippi Writing Assessment (WA) in grades 4 and 7. Results showed that, although program effects were not as strong and consistent as in the prior school year (Ross & Nunnery, 2004), they remained generally positive: 16 out of 20 computed effect size estimates were positive. In both Reading and Language Arts, significant positive program effects were observed in 5th grade and 8th grade. In Mathematics, a strong, significant program effect was observed in 3rd grade (d = +0.24), and a significant positive effect was also observed in 6th grade (d = +0.12). As in 2002-03, strong, significant program effects were obtained in both tested grade levels, particularly in 4th grade (ES = +0.52), on the Writing Assessment.

Supplementary analyses examined outcomes separately by school district, and examined the relationship of student-level implementation measures to achievement effects. Results showed much stronger and consistent program effects in Reading, Language Arts, Math, and Writing for Biloxi than for Pascagoula. Further, implementation measures (particularly reading level and comprehension scores) accounted for substantial variance in achievement, especially in lower grades. These results overall indicate generally positive and sustainable impacts of SR on student performance in program-independent, high-stakes testing. However, these impacts were predominantly evident in Biloxi and only minimally so in Pascagoula, thus suggesting that sustained implementation (four years vs. two years, respectively) was critical to success.

Introduction

The purpose of the present study was to compare the achievement of students in 14 Pascagoula schools and 9 Biloxi schools implementing School Renaissance (SR) to that of students in matched Control schools. The achievement measures to be examined were from the 2004 administration of the Mississippi Curriculum Test (MCT) for Reading, Language Arts, and Mathematics in grades 2-8 and the Mississippi Writing Assessment (WA) in grades 4 and 7. The Pascagoula schools began implementing SR in January, 2002, and the Biloxi schools began implementation in August, 2000. Thus, the present analyses of 2004 achievement data relates to the second and fourth full years of SR in Pascagoula and Biloxi, respectively. This report represents the second consecutive year in which SR program effects have been examined in these schools.

Summary of First Year Evaluation

SR effects on student achievement and school climate previously were examined for the 2002-03 school year by Ross and Nunnery (2004). SR teachers were significantly more favorable than Control teachers on all School Climate Inventory (SCI) dimensions, with a strong median effect size of +0.45. The dimensions reflecting the largest SR advantages were Collaboration, Environment, and Leadership. For both Reading and Language, SR students scored significantly higher than comparison students in 2003 after controlling for 2002 MCT scores. Program effects, while small, were generally consistent and positive across grades 3 to 8. Median effect size estimates across grades were +0.11 for Reading and +0.12 for Language. In Writing, fourth grade SR students scored significantly and substantially higher than comparison students, with an effect size estimate of +0.45. Nearly double the percentage of fourth grade SR students (44.5%) scored 3 or 4 on the Writing Assessment relative to the comparison group

(25.4%). No significant effects were observed for seventh grade Writing. Mathematics outcomes were significantly and substantially higher for SR students in grades 6 and 7, with effect size estimates of +0.22 and +0.27, respectively. No differences were observed in other grades. Achievement effects were stronger for SR schools in Biloxi (after three years of SR) than in Pascagoula (after only one year).

Research Questions

Given the promising results of the first year evaluation, the current report examines the impact of SR on student achievement in the 2003-04 school year, and whether the impact differs among students classified as low, medium, or high on prior measures of achievement. Specifically, the following research questions were addressed:

1. After controlling for 2003 achievement, is there is an overall difference between SR and comparison students on the Reading, Language Arts, Mathematics, or Writing subtests administered as part of the Mississippi school accountability program?
2. Is there an interaction between the program and students' prior levels of achievement; i.e., is the program more effective with low, middle, or high achievers?
3. Do results differ for a district having four years of implementation (Biloxi) vs. a district having only two years of implementation (Pascagoula)?
4. Are AR/RR results related to student implementation variables after adjusting for prior achievement and student characteristics?

Methodology

Participants and Design

A matched program-control school quasi-experimental design was employed for the current study. Participants included 8,264 students in grades 3 through 8 attending one of 41 schools in 15 school districts in southern Mississippi. The majority (60.1%) of the students were eligible for free or reduced-price lunch, nearly equal proportions were male (50.9%) or female (49.9%), and 43.7% were minority students (40.0% of the total were African American). A listing of matched School Renaissance (SR) and Comparison (C) school pairs is provided along with school demographic data for elementary and middle/high school samples in Table 1. In all, there were 23 SR schools and 18 comparison schools. Grade level configurations of treatment schools were seven K-6, eleven K-5, three 6-8, one 8-9, and one 7th grade only (see Table 1). Of the 18 comparison schools, three were PK-3, four were K-6, one was K-8, six were K-12, and four were middle schools (4-8).

A matched comparison school was selected for each SR school using the following criteria:

- the percentage of African American students enrolled for 2000-02;
- the percentage of White students enrolled for 2000-02;
- the percentage of economically-disadvantaged students enrolled for 2000-02;
- the percentage of mobile students as determined from 2000-02 cumulative attendance;
- School location (rural, suburban, small city);
- Grades served;
- School size;
- School means on 2001-02 MCT in Reading and Mathematics; and
- No or very limited usage of Accelerated Reader or Accelerated Math.

Table 1

School demographics for SR and matched Comparison schools[1]

School	Grade Level (Per CCD)	% Free or Reduced Lunch	% Monthly	Grade 3 %Prof&Adv Reading	Grade 5 %Prof&Adv Reading	Student/ Teacher Ratio
Rawls Springs Att Ctr	K-6	64.02	48.17	90.0	77.7	12.7
Popp's Ferry Elem	K-6	63.39	50.30	94.9	90.7	16.1
Dixie Elem School	K-8	64.18	24.55	71.7	59.1	14.2
North Bay Elem	K-6	63.40	28.11	94.6	97.2	17.1
Wheeler School (K-12)	K-12	63.64	11.22	86.1	56.3	13.2
Lopez Elem	K-6	63.40	46.39	96.6	97.1	17.9
Dixie Elem School	K-6	64.18	24.55	71.7	59.1	14.2
Jeff Davis Elem	K-6	63.46	42.12	92.5	91.7	14.6
Columbia Primary School	PK-3	67.31	49.84	91.3	NA	15.2
Gorenflo Elem	K-6	63.31	66.91	89.4	95.6	12.3
Scott Central Att Ctr	K-12	72.60	50.55	55.0	35.6	18.4
DuKate Elem	K-6	63.24	81.61	77.4	96.0	11.3
Wheeler School (K-12)	K-12	63.64	11.22	86.1	56.3	13.2
Beauvoir Elem	K-6	63.40	34.53	93.3	96.4	13.6
Morton Elementary School	PK-3	72.67	50.16	83.8	NA	16.5
Singing River	K-5	60.79	56.85	81.7	91.2	19.4
Perkinston Elem School	K-6	56.05	20.97	95.2	81.4	16.5
Martin Bluff Elem	K-5	60.77	23.60	70.7	95.0	18.5
North Pike Elem School	K-6	57.30	31.32	88.3	NA	19.6
Lake Elem	K-5	60.74	43.70	93.8	84.6	14.8
North Pike Elem School	K-6	57.30	31.32	88.3	NA	19.6
Jackson Elem	K-5	60.62	67.69	83.3	69.4	18.3
Kosciusko Upper Elem	K-6	59.88	55.02	NA	90.9	15.8
Gautier Elem	K-5	60.80	56.81	78.4	87.2	16.9
Second Street Elem	4-8/7-8	64.99	24.44	NA	92.0	17.1
Eastlawn	K-5	60.84	34.27	87.6	92.6	17.4
Thrasher High School	K-12	63.55	20.05	71.5	37.9	12.5
College Park	K-5	60.76	48.94	82.5	90.6	16.1
Lake Attendance Ctr	K-12	72.77	50.00	90.5	NA	18.1
Cherokee Elem	K-5	60.75	50.81	88.0	86.1	15.2
Edinburg Attendance Ctr	K-12	72.84	13.17	57.1	70.6	15.0
Central Elem	K-5	65.64	72.84	64.5	81.4	13.2
Walnut Attendance Ctr	K-12	65.21	17.72	71.4	68.2	14.9
Beach Elem	K-5	61.21	30.17	100.0	100.0	14.5
North Bay Elem School	PK-3	63.40	28.11	94.6	97.2	17.1
Arlington Heights	K-5	60.63	68.31	82.1	90.9	19.1
Byram Middle School	4-8/7-8	54.02	36.85	83.4	62.1	18.7
Biloxi Jr. High	8,9	56.14	56.77	NA	71.0	NA

Table 1

School demographics for SR and matched Comparison schools[1]

School	Grade Level (Per CCD)	% Free or Reduced Lunch	% Monthly	Grade 3 %Prof&Adv Reading	Grade 5 %Prof&Adv Reading	Student/ Teacher Ratio
Middle Sch of Poplarville	4-8/7-8	60.12	11.90	80.2	64.9	17.2
Michel 7th Grade	7	61.94	44.64	86.4	NA	NA
Walnut Attendance Center	K-12	65.21	17.72	71.4	68.2	14.9
Trent Lott Middle	6,7,8	60.83	36.28	80.6	77.0	17.3
Solomon Jr. High School	4-8/7-8	89.27	97.13	40.2	35.4	14.0
Colmer Middle	6,7,8	60.70	49.81	72.2	60.3	15.9
Byram Middle School	4-8/7-8	54.02	36.85	83.4	62.1	18.7
Gautier Middle	6,7,8	60.80	44.84	68.3	62.8	14.8

[1]The comparison school is listed first in each shaded or non-shaded pair.

Candidate comparison schools were identified through an examination of state-wide school data on the above variables. Following compilation of an initial listing of the strongest matches, the district leaders (superintendent or designee) associated with each of the candidate schools were first contacted by the researchers to determine interest and gain support. Comparison (C) schools would not be required to participate in any active way except to give permission for the researchers to obtain test scores from the MS DOE. In return for their cooperation, the C schools received $500 in cash. Altogether, two out of the original 15 district superintendents contacted declined. Those C schools were replaced by the next best available matches from other districts. All individual comparison schools on the list agreed, through their principals, to participate.

Measures

Student achievement. Mississippi Curriculum Test (MCT) subtests in Reading, Language Arts, and Mathematics, and the Mississippi Writing Assessment (WA), were used to measure student achievement. Three forms of the MCT are used in each subject area in grades 2 through 8. The Mississippi State Department of Education reported Cronbach's alpha reliability

coefficients across forms and grade levels ranging between 0.88 and 0.90 for Reading, 0.87 and 0.91 for Language, and 0.85 to 0.90 for Mathematics in 2001 (MSDOE, 2002). The MCT provides vertically-equated scale scores that use the same metric for student performance across grade levels, allowing for comparison of gains across grade levels and tracking of individual growth patterns (Tomkowicz & Schaeffer, 2002). Content validity of the MCT's was addressed by statewide teacher committees, who formed consensus about what specific skills and objectives were taught in particular subjects and grade levels (MSDOE, 2002). The WA is administered in grades 4 and 7, and yields an integer score between 0 and 4. Rubrics for scoring each writing assessment are provided in the appendix.

Procedure

Student records from participating schools for 2002-03 MCT, 2003-04 MCT, and 2003-04 WA were provided by the Mississippi State Department of Education. Records were matched by student identification number. Students were included in the analysis only if they had scores on all three MCT subtests for both 2002-03 and 2003-04.

Analysis

To allow examination of the interaction between program and prior student performance level, for each MCT subtest and for the WA, students were divided into three groups based upon performance on that subtest during the prior year: low ability (below 33rd percentile), middle ability (between 33rd and 66th percentile), and high ability (above 66th percentile). A 2 (SR versus C) X 3 (ability group) analysis of covariance was performed for each grade level on the Reading, Language Arts, and Mathematics subtests of the MCT, and for 4th and 7th grades on the WA. Cohen's *d* was computed as an effect size measure for each comparison by subtracting the

comparison group adjusted mean from the treatment group adjusted mean, and dividing by the comparison group standard deviation.

Results

Combined District Analyses

Reading

3rd grade. ANCOVA indicated no significant treatment effects ($F_{1,1462}$=1.709, p=0.19) and no treatment by ability group interaction effect ($F_{2,1462}$=2.62, p=0.07). Results directionally favored SR students (*M*'=509.2) over comparison students (*M*'=507.1; see Table 2). The effect size estimate was d = +0.05.

4th grade. As with third grade, no significant treatment ($F_{1,1399}$=0.18, p=0.62) or treatment by ability group interaction effect ($F_{2,1399}$=0.20, p=0.82) was observed. Covariate-adjusted posttest means were nearly equal for SR (*M*'=519.0) and comparison students (*M*'=519.6; see Table 2). The effect size estimate was d = -0.01.

5th grade. A significant treatment effect was observed in fifth grade ($F_{1,1297}$=5.26, p=0.02). The interaction effect was not significant ($F_{2,1297}$=0.24, p=0.79), indicating the treatment effect was constant across levels of prior achievement. SR students (*M*'=542.2) scored significantly higher than comparison students (*M*'=538.6) after controlling for pretest scores (see Table 2). The effect size estimate was d = +0.09.

6th grade. No significant treatment ($F_{1,1062}$=0.47, p=0.50) or interaction effects ($F_{2,1062}$=0.86, p=0.43) were observed. The pretest adjusted posttest mean for comparison students (*M*'=546.1) was slightly higher than that for SR students (*M*'=544.6; see Table 2). The effect size estimate was d = -0.04.

7th grade. The ANCOVA showed no significant differences related to treatment condition ($F_{1,1394}$=1.33, p=0.25) or the treatment by ability group interaction ($F_{2,1394}$=0.04, p=0.96). SR students had a marginally higher adjusted posttest mean (M'=559.8) than comparison students (M'=558.1; see Table 2). The effect size estimate was d = +0.04.

8th grade. A significant treatment effect was observed for eighth grade reading ($F_{1,1608}$=8.54, p=0.004), indicating the SR students scored significantly higher than comparison students after controlling for pretest scores. Adjusted means were M'=578.2 for SR students, and M'= 574.0 for comparison students (see Table 2). The effect size estimate was d = +0.10.

Table 2

2004 MCT Reading Means by Grade and Treatment Condition

Grade/Treatment	*M*	*M'*	*SD*	*n*	*ES*
Third					
Control	504.5	507.1	45.0	689	+0.05
SR	511.7	509.2	41.8	780	
Fourth					
Control	519.0	519.6	41.1	615	-0.01
SR	520.8	519.0	41.1	791	
Fifth					
Control	537.6	538.6	40.6	556	+0.09
SR	543.3	542.2*	38.4	748	
Sixth					
Control	541.9	546.1	37.1	258	-0.04
SR	546.4	544.6	43.5	811	
Seventh					
Control	556.6	558.1	42.2	644	+0.04
SR	561.5	559.8	42.3	757	
Eighth					
Control	568.2	574.0	42.6	764	+0.10
SR	583.9	578.2*	42.1	851	

*Significantly higher than comparison group mean at $p < .05$.

Language Arts

3rd grade. No treatment ($F_{1,1462}$=0.34, p=0.56) or interaction effects ($F_{2,1462}$=0.63, p=0.53) were observed. SR students had directionally higher adjusted posttest means (*M*'=516.6) than comparison students (*M*'=515.4; see Table 3). The effect size estimate was d = +0.02

4th grade. ANCOVA indicated no significant treatment ($F_{1,1399}$=0.17, p=0.69) or interaction effects ($F_{2,1399}$=0.42, p=0.66). SR students had a slightly higher adjusted posttest mean (*M*'=527.2) than comparison students (*M*'=526.4; see Table 3). The effect size estimate was d = +0.02.

5th grade. A significant treatment effect ($F_{1,1297}$=16.17, p<.001) was observed for 5th grade Language Arts scores. The treatment by ability group interaction effect was not significant ($F_{2,1297}$=2.37, p=0.09). The adjusted mean for SR students (*M*'=551.3) was significantly higher than the adjusted mean for comparison students (*M*'=545.2; see Table 3). The effect size estimate was d = +0.14.

6th grade. No significant treatment ($F_{1,1062}$=1.65, p=0.20) or interaction ($F_{2,1062}$ = 0.39, p=0.68) effects were indicated. The adjusted mean for the comparison group (*M*'=551.1) was somewhat higher than the adjusted mean for SR students (*M*'=548.4; see Table 3). The effect size estimate was d = -0.06.

7th grade. The analyses showed no significant differences between SR and comparison students on 2004 Language Arts scores ($F_{1,1394}$ = 1.55, p=0.21), and no treatment by ability group interaction effect ($F_{2,1394}$=2.81, p=0.06). The adjusted mean for SR students (*M*'=573.7) was directionally higher than the adjusted mean for comparison students (*M*'=571.4; see Table 3). The effect size estimate was d = +0.04.

8th grade. As shown in Table 3, SR eighth-grade students had a significantly higher adjusted posttest mean on Language Arts (*M*'=572.2) than comparison students (*M*'=569.2; $F_{1,1608}$=4.03, p<.05). The effect size estimate was d = +0.07.

Table 3

2004 MCT Language Arts Means by Grade and Treatment Condition

Grade/Treatment	*M*	*M'*	*SD*	*n*	*ES*
Third					
Control	514.9	515.4	49.2	689	+0.02
SR	519.4	516.6	50.7	780	
Fourth					
Control	526.3	526.4	50.0	615	+0.02
SR	528.8	527.2	48.6	791	
Fifth					
Control	544.7	545.2	42.2	556	+0.14
SR	552.7	551.3*	40.1	748	
Sixth					
Control	546.0	551.1	42.2	258	-0.06
SR	550.4	548.4	45.6	811	
Seventh					
Control	568.7	571.4	54.5	644	+0.04
SR	576.6	573.7	51.8	757	
Eighth					
Control	565.0	569.2	44.7	764	+0.07
SR	577.1	572.2*	45.8	851	

*Significantly higher than comparison group mean at p < .05.

Mathematics

3rd grade. ANCOVA indicated a significant treatment effect ($F_{1,1462}$=26.12, p<0.01) but no treatment by ability group interaction effect ($F_{2,1462}$=2.66, p=0.07). SR students had a significantly higher adjusted mean (*M*'=516.2) than comparison students (*M*'=505.2; see Table 4). The effect size estimate was d = +0.24.

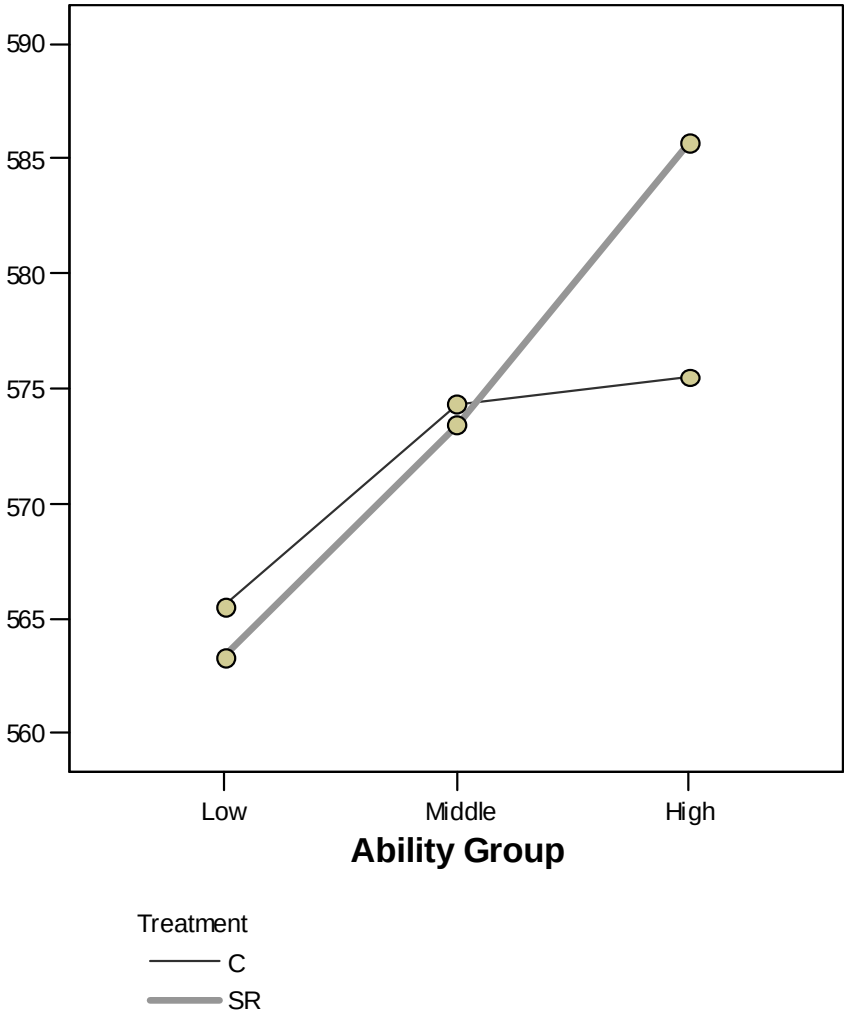
4th grade. No significant treatment ($F_{1,1399}$=0.01, p=0.62) or treatment by ability group interaction effect ($F_{2,1399}$=2.66, p=0.12) was observed. Covariate-adjusted posttest means were equal for SR (M'=530.5) and comparison students (M'=530.5; see Table 4). The effect size estimate was d = 0.00.

5th grade. A significant treatment effect was observed in fifth grade ($F_{1,1297}$=9.95, p=0.002). The interaction effect was not significant ($F_{2,1297}$=1.13, p=0.33), indicating the treatment effect was constant across levels of prior achievement. SR students (M'=561.2) scored significantly higher than comparison students (M'=555.7) after controlling for pretest scores (see Table 4). The effect size estimate was d = +0.12.

6th grade. No significant treatment effect ($F_{1,1062}$=1.27, p=0.26) was observed, but the treatment by ability group interaction effect was significant ($F_{2,1062}$=6.58, p=0.001). The overall adjusted posttest mean for SR students (M'=574.1) was slightly higher than that for comparison students (M'=571.8; see Table 4). Follow-up analyses showed that SR students in the high ability group scored significantly higher (M'=585.7) than comparison students in the high group (M'=575.5), while treatment group means did not differ in the low and middle ability groups (see Figure 1). The overall effect size estimate was d = +0.06, while the high-ability group effect size estimate was d = +0.26.

Figure 1

Interaction between Treatment Condition and Ability Group: Sixth Grade



Note: SR adjusted mean at high ability grouping is significantly higher than C mean at $p<.05$.

7th grade. The ANCOVA showed no significant differences related to treatment condition ($F_{1,1394}=3.11$, $p=0.08$) or treatment by ability group interaction ($F_{2,1394}=0.71$, $p=0.49$). SR students had a marginally higher adjusted posttest mean ($M'=586.4$) than comparison students ($M'=583.6$; see Table 4). The effect size estimate was $d = +0.06$.

8th grade. The treatment effect for eighth grade reading was not significant ($F_{1,1608}$=1.95, p=0.16), but there was a significant treatment by ability group interaction effect ($F_{1,1608}$=4.23, p=0.015). Adjusted overall means were *M*'=600.8 for SR students, and *M*'= 602.6 for comparison students (see Table 4). Follow-up analyses showed that comparison students in the high ability group had a significantly higher adjusted mean (*M*'=620.4) than SR students in the high ability group (*M*'=615.0), while means between the treatment groups were equal in the low and middle ability groups (see Figure 2). The overall effect size estimate was *d* = -0.04, while the high ability effect size estimate was *d* = -0.13.

Table 4

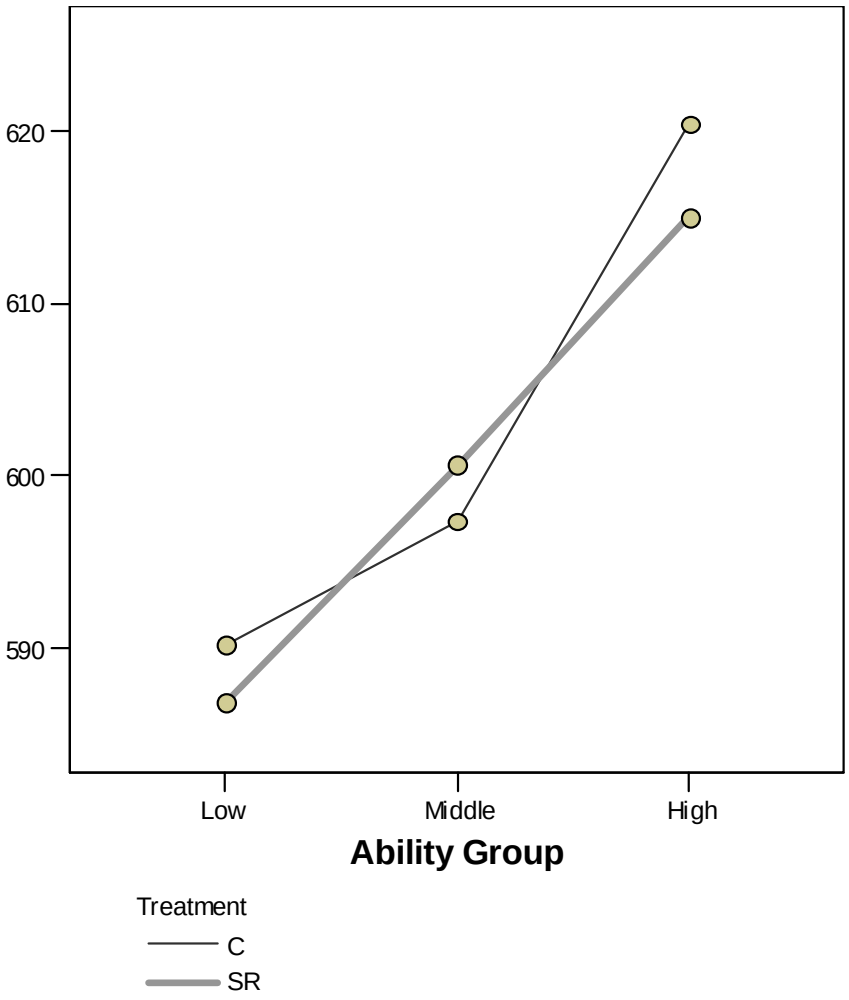
2004 MCT Mathematics Means by Grade and Treatment Condition

Grade/Treatment	*M*	*M'*	*SD*	*n*	*ES*
Third					
Control	508.0	505.2	46.3	689	+0.24
SR	514.5	516.2*	49.4	780	
Fourth					
Control	528.0	530.5	43.1	615	+0.00
SR	532.6	530.5	43.5	791	
Fifth					
Control	555.8	555.7	44.7	556	+0.12
SR	561.9	561.2*	48.3	748	
Sixth					
Control	570.3	571.8	39.0	258	+0.06
SR	575.7	574.1[a]	47.3	811	
Seventh					
Control	581.8	583.6	46.0	644	+0.06
SR	589.2	586.4	49.7	757	
Eighth					
Control	594.3	602.6	40.6	764	-0.04
SR	609.0	600.8[a]	43.7	851	

[a]Significant treatment by ability group interaction effect (see Figures 1 and 2).
*Significantly higher than comparison group mean at p<.05

Figure 2

Interaction between Treatment Condition and Ability Group: Eighth Grade



Note: C adjusted mean at the high ability grouping significantly higher than SR mean at $p<.05$.

Writing

4th grade. ANCOVA indicated a strong and significant program effect on 2004 Writing scores ($F_{1,1394}$=89.5, p<.001), with SR students having a significantly higher adjusted mean (M'= 2.77) than comparison students (M' = 2.40; see Table 5). The treatment by ability group interaction effect was not significant ($F_{2,1394}$=1.61, p=0.20). The effect size estimate was d = +0.52.

7th grade. A significant treatment effect was observed for 2004 Writing scores in 7th grade ($F_{1,1371}$=3.92, p<.05). The SR adjusted posttest mean (M' = 2.76) was significantly higher than the comparison adjusted posttest mean (M'= 2.68; see Table 5). There was no significant treatment by ability group interaction ($F_{2,1371}$=2.31, p=0.10). The effect size estimate was d = 0.12.

Table 5

2004 MCT Writing Means by Grade and Treatment Condition

Grade/Treatment	*M*	*M'*	*SD*	*n*	*ES*
Fourth					
Control	2.40	2.40	0.71	612	+0.52
SR	2.78	2.77*	0.78	789	
Seventh					
Control	2.67	2.68	0.69	638	+0.12
SR	2.79	2.76*	0.79	740	

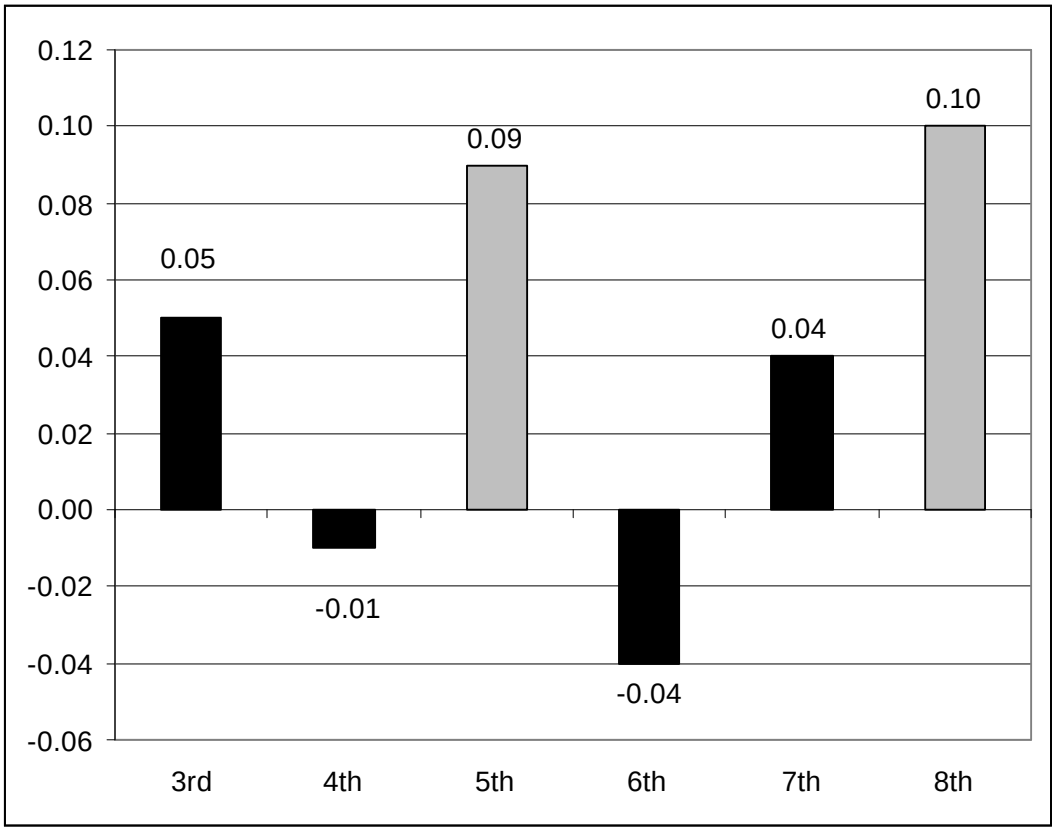
*Significantly higher than comparison group mean at p<.05.

Effect Size Summary

Reading and Language Arts effects showed the same pattern of generally small effect size estimates (Reading d = -0.04 to +0.10; Language d = -0.06 to +0.14), with modest but significant effect sizes in grades 5 and 8 for both subject areas (see Figures 3 and 4). Mathematics effect size estimates ranged from small (d = -0.04) to moderately large (d = +0.24), with significant overall effects in 3rd grade and 5th grade (see Figure 5). In Writing, program effects were substantial and significant: d for 4th grade was +0.52, and d for 7th grade was +0.12 (see Figure 6).

Figure 3

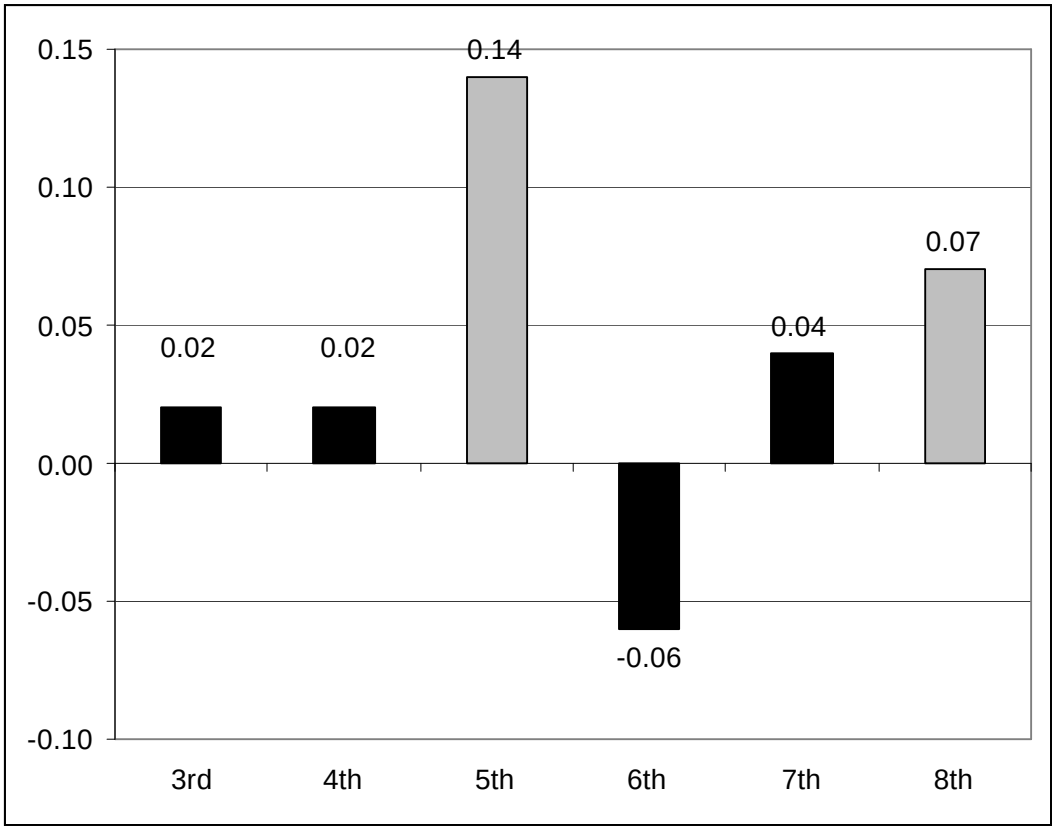
Reading Effect Size Estimates by Grade Level



Note: Light gray bars indicate significant effects.

Figure 4

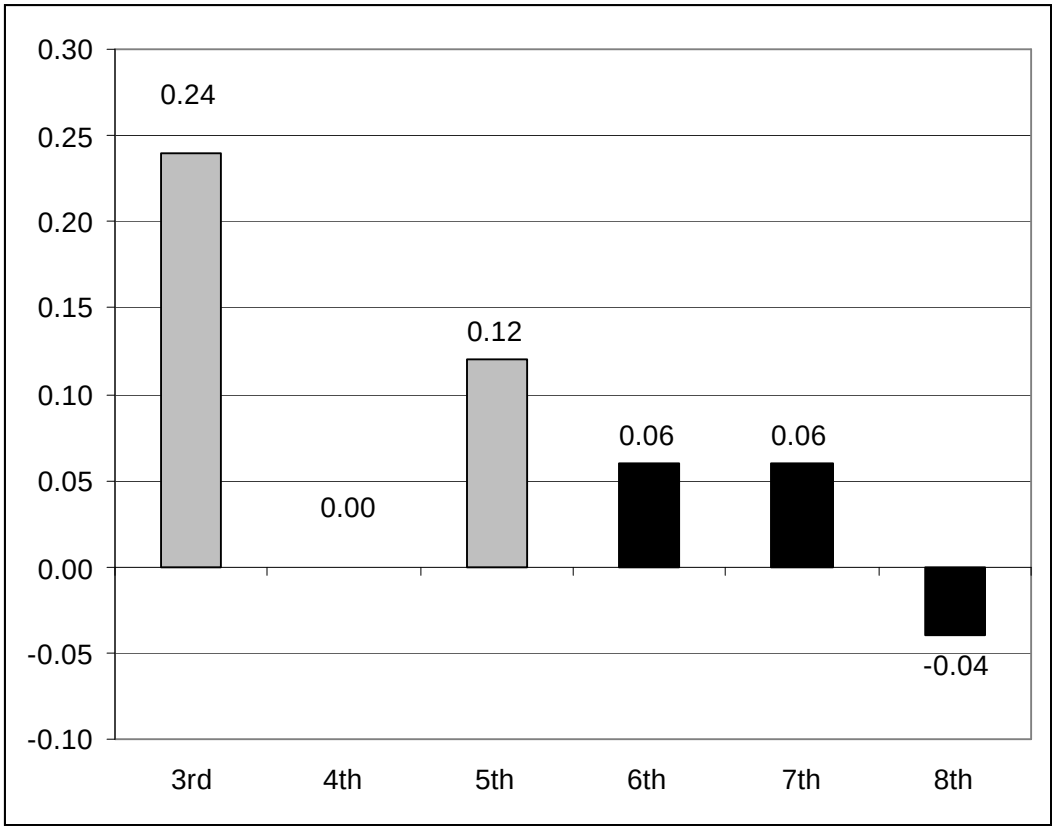
Language Arts Effect Size Estimates by Grade Level



Note: Light gray bars indicate significant effects.

Figure 5

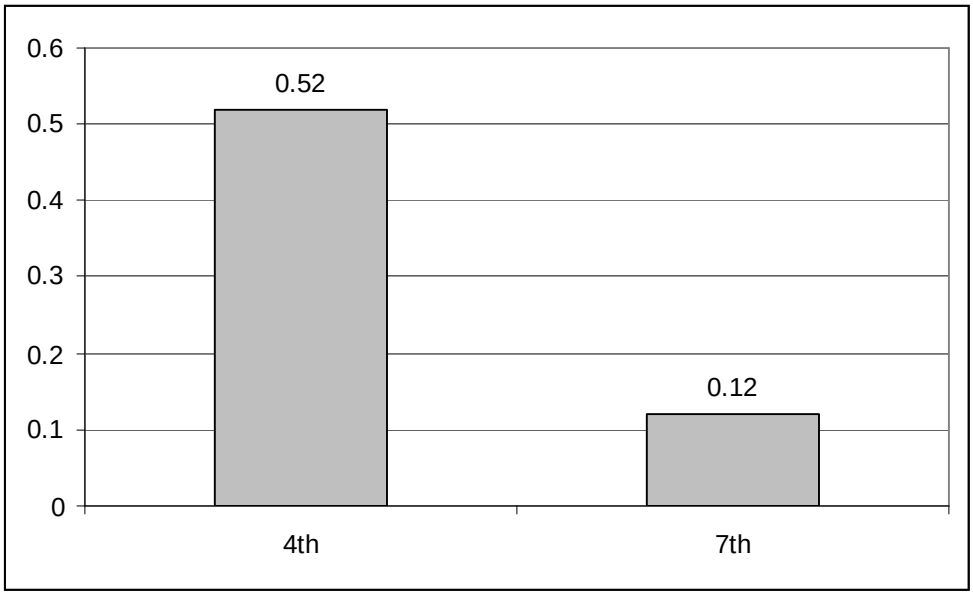
Mathematics Effect Size Estimates by Grade Level



Note: Light gray bars indicate significant effects.

Figure 6

Writing Assessment Effect Size Estimates by Grade Level



Note: Light gray bars indicate significant effects.

Analyses of District Effects

To determine whether length of experience with implementation of School Renaissance was associated with program effects on student achievement, exploratory analyses of covariance were conducted on 2004 MCT Reading, Language, Mathematics scores, and WA scores. District (Pascagoula, Biloxi, and Control) served as the independent variable, and corresponding 2003 MCT scale scores served as covariates. MCT Language was used as a covariate for the Writing Assessment analyses. Analyses were performed for each grade level. Effect size estimates were computed by dividing the difference between treatment district and Control schools adjusted means by the total standard deviation for Control group students. Pascagoula had been implementing SR for two years, and Biloxi for four years, at the time of 2004 MCT testing.

Reading: District Comparisons

In third grade, a significant difference ($F_{2,1465}$=4.64, p=0.01) was observed among districts. Follow-up tests indicated that the adjusted mean for Biloxi students (M'=513.1) was significantly higher than both the Control group adjusted mean (M'=507.3) and the Pascagoula adjusted mean (M'=506.6; see Table 6). Effect size estimates were -0.02 and +0.13 for Pascagoula and Biloxi, respectively (see Figure 7). A significant district effect was also observed for fourth grade ($F_{2,1402}$=3.84, p=0.02), with Biloxi students (M'=523.1; ES = +0.07) scoring significantly higher than Pascagoula students (M'=517.4; ES = -0.07; see Table 6 and Figure 7). Significant differences were also observed in 5th ($F_{2,1300}$ = 3.13, p <.05) and 6th grades ($F_{2,1065}$=17.19, p<.001), with Biloxi students scoring significantly higher than Control students in both grades, and higher than Pascagoula students in 6th grade. Control students also scored significantly higher than Pascagoula students in 6th grade. Effect size estimates for 5th and 6th grade in Biloxi were +0.11 and +0.15, respectively, compared to +0.09 and -0.18 in Pascagoula (see Figure 7). No significant differences were observed in 7th grade, but a significant difference across districts was observed for 8th grade ($F_{2,1611}$=6.52, p<.01). Follow-up tests indicated that Pascagoula students (M'=580.2, ES = +0.15) scored higher than Control students (see Figure 7).

Language Arts

Significant differences were observed at all grade levels except grade 3. In grade 4, Biloxi students (M'=533.4; ES = +0.12) scored significantly higher ($F_{2,1402}$=6.15, p<.01) than both Control (M' = 527.4) and Pascagoula students (M' = 523.8; ES = -0.07; see Table 6 and Figure 8). In fifth grade, a significant difference ($F_{2,1300}$=7.65, p<.001) was found favoring both Pascagoula (M'=552.6, ES = +0.16) and Biloxi (M'=551.0; ES = +0.12) over Control students. At the sixth grade level, Biloxi students (M'=555.9; ES = +0.11) scored significantly higher

($F_{2,1065}$=19.55, p <.001) than both Pascagoula (*M*' = 543.4; *ES* = -0.19) and Control students (*M*'=551.2). A significant difference was also observed in seventh grade ($F_{2,1397}$=3.28, p<.05). Follow-up tests showed that Biloxi students (*M*'=577.5; *ES* = +0.10) scored significantly higher than both Pascagoula (*M*'=571.5; *ES* = -0.01) and Control students (*M*'=571.8). Finally, a significant difference in eighth grade ($F_{2,1611}$=3.30, p<.05) was attributable to Pascagoula students (*M*'=574.0; *ES* = +0.10) scoring higher than Control students (*M*'=569.5; see Figure 8).

Mathematics

Although no differences were observed in 7th or 8th grades, significant differences among districts were observed in 3rd grade ($F_{2,1465}$=11.19, p<.001), 4th grade ($F_{2,1402}$=31.26, p<.001), 5th grade ($F_{2,1300}$=6.19, p<.01), and 6th grade ($F_{2,1065}$=26.59, p<.001). In third grade, both Pascagoula (*M*'=516.5; *ES* = +0.22) and Biloxi (*M*' = 515.9; *ES* = +0.21) students scored significantly higher than Control students (*M*'=506.1; see Table 6 and Figure 9). Fourth grade Biloxi students (*M*' = 540.5; *ES* = +0.23) scored higher than both Pascagoula (*M*' = 522.8; *ES* = -0.18) or Control students (*M*' = 530.7). In fifth grade, both Pascagoula (*M*'=561.4; *ES* = +0.13) and Biloxi (*M*' = 562.5; *ES* = +0.15) students scored significantly higher than Control students (*M*'=555.8). Follow-up tests in sixth grade indicated that Biloxi students (*M*'=583.8; *ES* = +0.30) scored significantly higher than Pascagoula (*M*'=568.7; *ES* =-0.09) and Control students (*M*'=572.3).

Writing

Significant differences among districts were observed for both fourth grade ($F_{2,1397}$=56.64, p <.001) and seventh grade ($F_{2,1397}$=85.49, p<.001) Writing Assessment scores. In fourth grade, both Pascagoula (M' = 2.88; *ES* = +0.66) and Biloxi (M' = 2.65; *ES* = +0.34) students scored significantly higher than Control students (M'=2.41; see Table 6 and Figure 10). In seventh grade,

Biloxi students (M'=3.15; *ES* = +0.67) scored significantly higher than both Pascagoula (M'=2.51; *ES* = -0.26) and Control students (M'=2.69). Seventh grade Control students also scored significantly higher than Pascagoula students.

Table 6

2004 Reading, Language, Mathematics, and Writing Adjusted Means by Grade Level and District

	Reading			Language			Mathematics			Writing		
	C	P	B	C	P	B	C	P	B	C	P	B
3rd Grade	507.3	506.6	513.1 [1,2]	516.9	517.8	517.6	506.1	516.5[1]	515.9[1]			
4th Grade	520.1	517.4	523.1[2]	527.4	523.8	533.4[1,2]	530.7[2]	522.8	540.5[1,2]	2.41	2.88[1,3]	2.65[1]
5th Grade	538.6	542.2	543.0[1]	546.0	552.6[1]	551.0[1]	555.8	561.4[1]	562.5[1]			
6th Grade	546.4[2]	539.9	552.0[1,2]	551.2[2]	543.4	555.9[1,2]	572.3	568.7	583.8[1,2]			
7th Grade	558.4	559.0	561.3	571.8	571.5	577.5[1,2]	584.0	586.9	588.0	2.69[2]	2.51	3.15[1,2]
8th Grade	574.0	580.2[1]	577.0	569.5	574.0[1]	572.1	602.6	602.0	600.9			

[1]Significantly higher than comparison district mean.
[2]Significantly higher than Pascagoula mean.
[3]Significantly higher than Biloxi mean.
Note. Schools from multiple districts were included in the comparison group means. Pascagoula had implemented School Renaissance for two years, and Biloxi for four years, at the time of the study.

Figure 7

2004 MCT Reading Effect Size Estimates by District and Grade Level

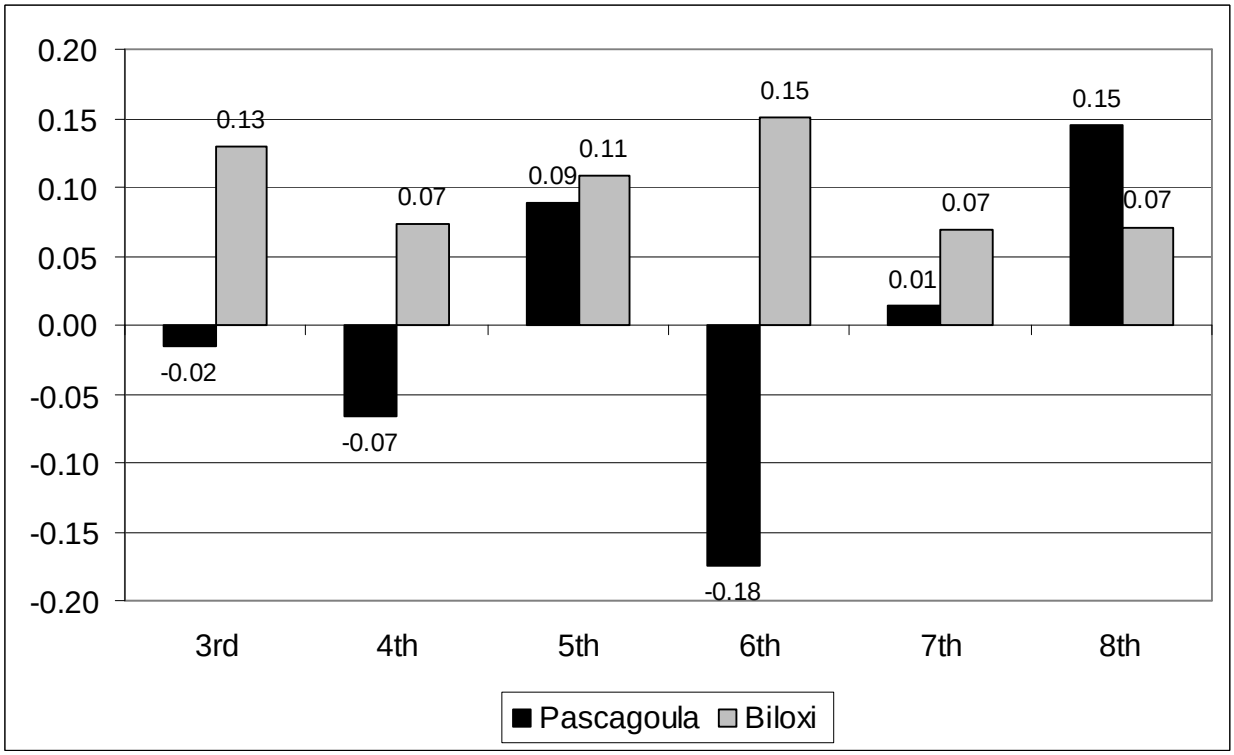


Figure 8

2004 MCT Language Effect Size Estimates by District and Grade Level

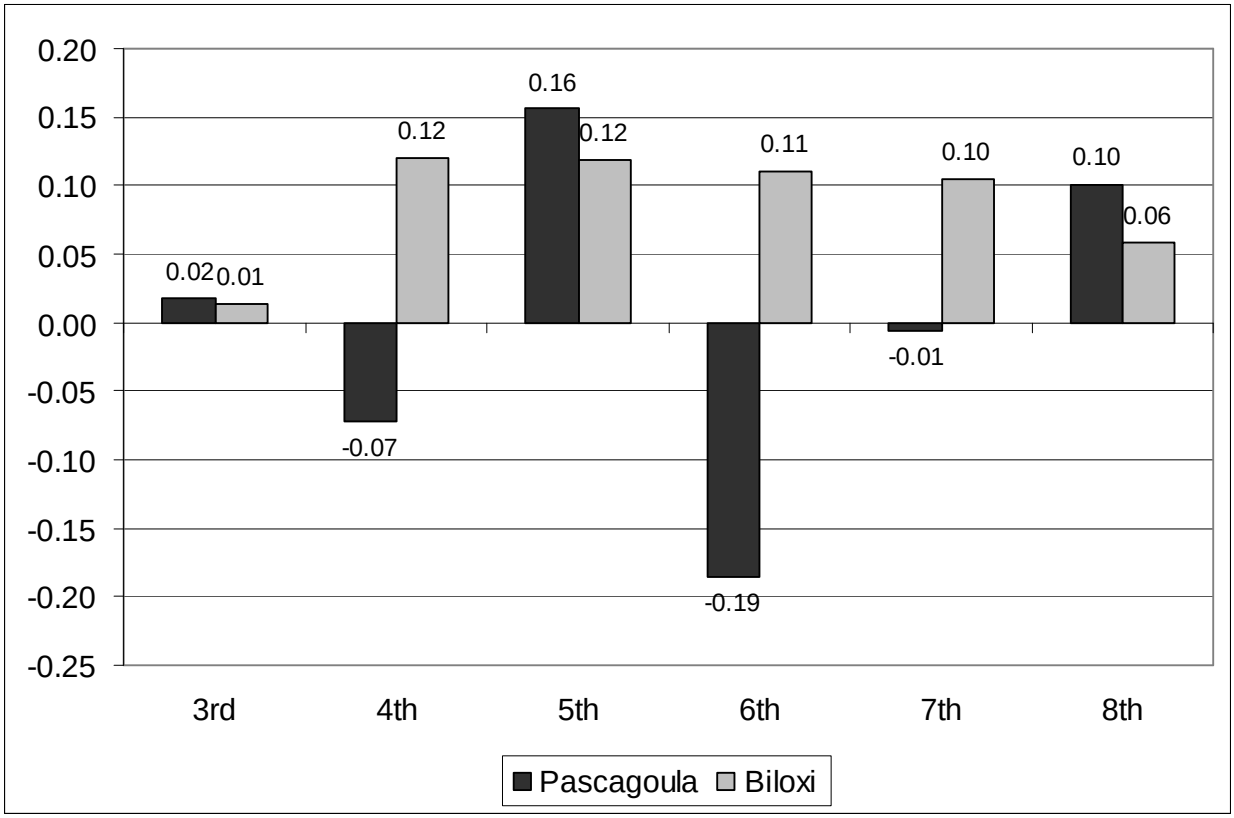


Figure 9

2004 MCT Mathematics Effect Size Estimates by District and Grade Level

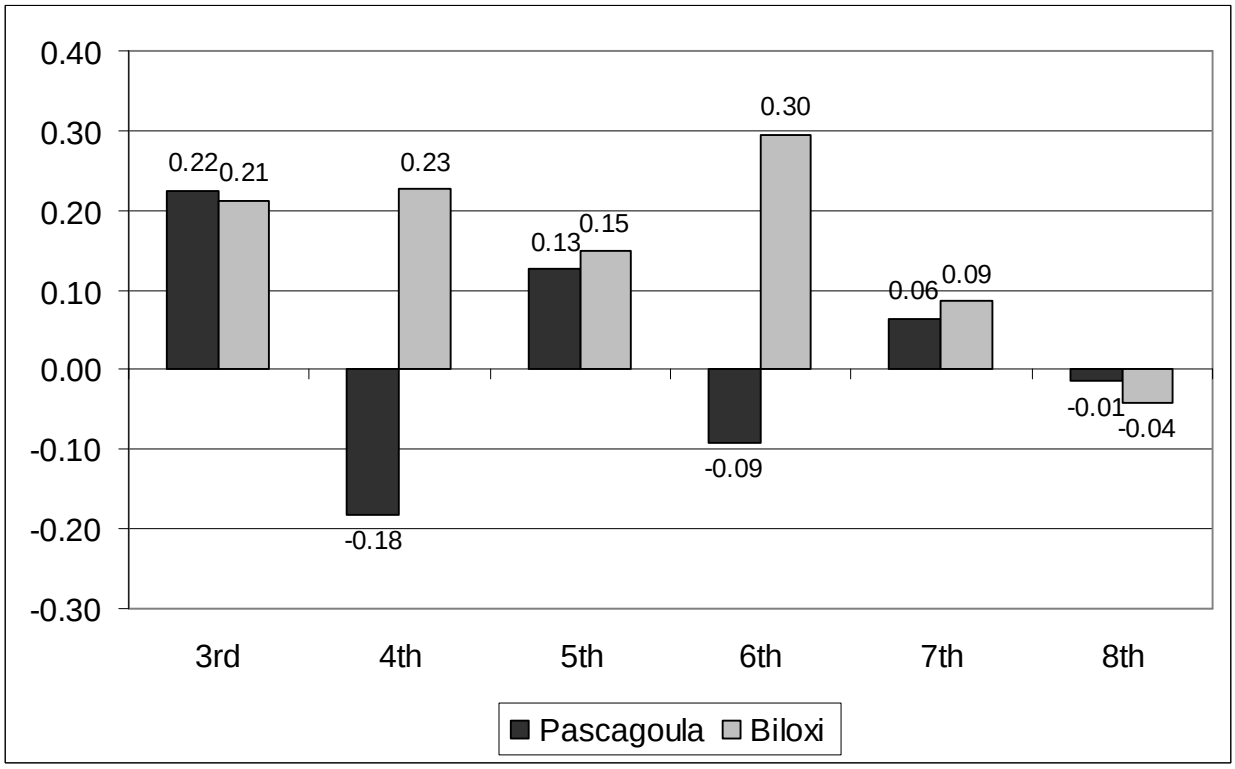
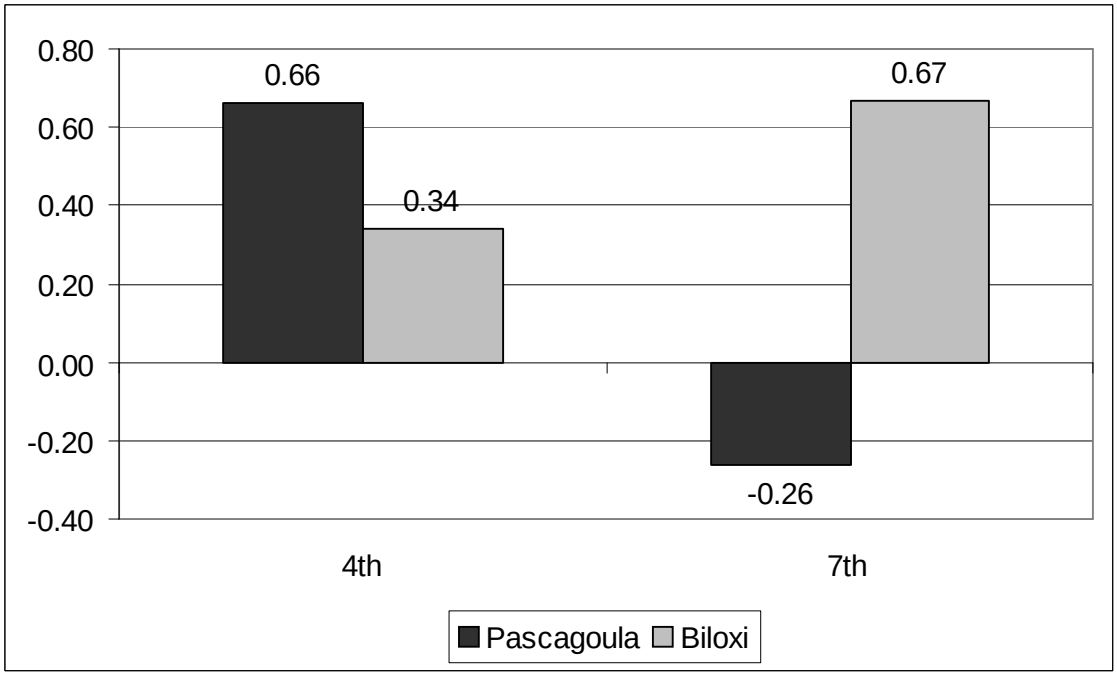


Figure 10

2004 Writing Assessment Effect Size Estimates by District and Grade Level



Relationship between Implementation and Achievement

Multiple regression analyses were used to determine the relationship between implementation and achievement gains in Reading and Mathematics. Implementation data were available at the individual student level. For Reading, four implementation variables were examined: (a) the number of books read during the school year (NUMBOOKS); (b) the number of points earned during the school year (POINTS); (c) average percent correct on quizzes (CORRECT); and (d) the average book level during the school year (LEVEL). A two-stage block entry multiple regression analysis was conducted for each grade level. The first block was comprised of 2003 achievement scores, free or reduced-price lunch status, gender, and special education status. The second block was comprised of the implementation variables. The *F*-test associated with the change in R^2 from the entry of the second block of predictor variables

indicates whether implementation predicts gains in achievement beyond that which is predictable by student background characteristics. Individual *t*-tests associated with the regression coefficients were performed to determine which specific implementation variables were significantly predictive of achievement gains after controlling for student background variables. Implementation analyses were not done for AM because insufficient cases with implementation data were available within grade levels to meet data requirements for regression modeling.

AR Implementation and Reading Achievement

Regression results indicated that AR implementation measures predicted substantial and significant variance in 2004 MCT Reading scores at every grade level, after controlling for 2003 MCT Reading scores, gender, special education status, and free or reduced-price lunch status. R^2 change generally decreased as grade level increased, from 22% of variance in 3rd grade, to about 10% in grades 4 through 6, to 5% in grades 7 and 8 (see Table 7). The most important predictor was *average reading level*, with standardized regression coefficients ranging from 0.19 to 0.45, all of which were significant at $p < .01$ (see Table 7). *Average percent correct* was also a consistently strong and significant predictor, with significant (at $p < .01$) standardized regression coefficients ranging from 0.11 to 0.27. *Number of points earned* was a significant predictor in 5th grade ($b = 0.07$, $p<.01$) and 7th grade ($b = 0.08$, $p<.01$), whereas *number of books read* was not a significant predictor at any grade level.

These results therefore suggest that mere volume of reading or book selection is not predictive of student achievement on standardized tests. Importantly, strongly predictive of such success, even after adjusting for prior achievement, ethnicity, gender, and poverty, is level of student attainment in using the AR/RR books and assessments.

Table 7

Accelerated Reader Implementation Regression Results

		Standardized Regression Coefficients				
Grade	R^2 change	NUMBOOKS	POINTS	CORRECT	LEVEL	*n*
3rd	0.22[1]	0.06	0.00	0.27[1]	0.45[1]	768
4th	0.09[1]	-0.02	0.02	0.19[1]	0.31[1]	772
5th	0.10[1]	0.02	0.07[1]	0.17[1]	0.33[1]	733
6th	0.10[1]	0.06	0.04	0.12[1]	0.36[1]	789
7th	0.05[1]	0.00	0.08[1]	0.11[1]	0.22[1]	728
8th	0.05[1]	0.00	0.03	0.16[1]	0.19[1]	820

[1]Significant at $p<.01$.
Note. R^2 change indicates the proportion of variance attributable to implementation variables after accounting for variance attributable to prior achievement, free or reduced-price lunch status, gender, and special education status.

Discussion and Conclusions

Combined District Outcomes

Although program effects were not as strong and consistent as in the 2002-03 school year, they remained generally positive: 16 out of 20 computed effect size estimates were positive. In both Reading and Language Arts, significant positive program effects were observed in 5th grade and 8th grade. In Mathematics, a strong, significant program effect was observed in 3rd grade (d = +0.24), and a significant positive effect was also observed in 6th grade (d = +0.12). As in 2002-03, strong, significant program effects were obtained in both tested grade levels, particularly in 4th grade (ES = +0.52), on the Writing Assessment.

In most subject areas, no treatment by ability group interaction effects were observed, indicating that the program effects were relatively constant across levels of students' prior achievement. Interaction effects were obtained, however, in 6th and 8th grades on the Mathematics subtest, with high achieving SR students performing significantly better than high achieving Control students in 6th grade, while high achieving Control students performed significantly better than high achieving SR students in 8th grade. Given the lack of consistency in the observed interactions, and the general lack of interaction overall, it appears that program effects are relatively constant across levels of prior achievement.

District and Implementation Outcomes

Comparisons by district showed that Biloxi students consistently achieved higher than students attending Control schools in every subject at every grade level. In 95% (19 of 20) of the comparisons performed in the study, students attending Biloxi SR schools scored directionally higher than students attending Control schools. In 75% of the comparisons, Biloxi students scored significantly higher than students attending Control schools. Exceptions to this pattern tended to occur in 7th and 8th grades, where Biloxi students scored comparably to Control students on all measures except 7th grade Writing Assessment. Effect size estimates indicated small to moderate effects in Reading and Language Arts, ranging from *ES* = +0.01 to +0.15, and moderately large to large effects in Mathematics (median *ES* = +0.22) and Writing (*ES* = +0.34 in fourth grade, and *ES* = +0.67 in 7th grade).

In contrast, Pascagoula students scored directionally higher than Control students on exactly half of the comparisons (10/20), and significantly higher on only 30% (6/20). Control students scored significantly higher than Pascagoula students on four comparisons. Effect size estimates tended to be small and inconsistent across grade levels and subject areas. These results

suggest the benefits of a longer program implementation period, in this case four vs. two years in Biloxi and Pascagoula, respectively. However, due to the confounding of school district characteristics with implementation time, this interpretation needs to be made cautiously.

AR implementation measures proved to be powerful and significant predictors of reading achievement, accounting for 5% to 22% of the variance in 2004 MCT Reading scores after controlling for prior achievement, free lunch status, gender, and special education status. The most important predictors were *average reading level* and *average percent correct*, which were both significant at all grade levels. *Number of points earned* was also a significant predictor in grades 5 and 7.

References

Mississippi Department of Education (2002). *Mississippi Curriculum Test: Summary of Technical Information.* Jackson, MS: Office of Research and Statistics, Mississippi Department of Education.

Ross, S.M., & Nunnery, J.A. (2004). *The effect of School Renaissance on student achievement in two Mississippi school districts.* Memphis, TN: Center for Research in Educational Policy.

Tomkowicz, J., & Schaeffer, G.A. (2002, April). *Vertical scaling for custom criterion-referenced tests.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans.

Appendix

Fourth Grade Mississippi Writing Assessment Scoring Rubric

SCORE 4

The student's writing

a. is about the topic (fully develops the writing prompt)
b. includes several details that support the topic
c. is organized (maintains logical sequence)
d. frequently contains interesting words (grade-level vocabulary or above)
e. contains complete sentences
f. follows punctuation, capitalization, spelling (both correct and phonetic), and usage rules (Two types of errors may occur: 1) those that appear as a consequence of risk-taking and 2) those that do not detract from overall quality.)

SCORE 3

The student's writing

a. is about the topic (partially develops the writing prompt)
b. includes some details that support the topic
c. is organized (may not maintain logical sequence throughout)
d. contains some interesting words (grade-level vocabulary)
e. contains complete sentences (may have occasional fragments and/or run-on sentences)
f. follows punctuation, capitalization, spelling (both correct and phonetic), and usage rules (Occasional errors occur that may detract from overall quality.)

SCORE 2

The student's writing

a. is about the topic (minimally develops the writing prompt)
b. includes only a few details that support the topic
c. shows minimal organization
d. contains only a few interesting words (grade-level vocabulary)
e. contains complete sentences (may have numerous fragments and/or run-on sentences)
f. rarely follows correct punctuation, capitalization, spelling, and usage rules

SCORE 1

The student's writing

a. attempts to address the topic (may digress from the writing prompt)
b. includes vague or no details that support the topic
c. shows no organization
d. includes no interesting words (below grade-level vocabulary)
e. contains numerous fragments and/or run-on sentences (may contain a complete sentence)
f. does not follow correct punctuation, capitalization, spelling, or usage rules

SCORE 0

The student's writing

a. is incomprehensible
b. is insufficient to score

Seventh Grade Mississippi Writing Assessment Scoring Rubric

SCORE 4

The student's response

a. addresses the specific writing prompt (fully develops the topic)
b. contains a clearly stated main idea (thesis)
c. shows a sense of audience and purpose
d. contains a minimum of three indented (or clearly delineated) paragraphs
e. has a clear beginning, middle, and end
f. has a main idea developed by supporting details that are well elaborated
g. exhibits logical order and appropriate sequencing of steps or ideas with adequate transitions
h. contains precise and vivid language (grade-level vocabulary or above)
i. maintains a consistent point of view
j. contains no errors in grammar usage that detract from the overall delivery (Grammar/usage includes subject-verb agreement, verb tense, pronoun case and reference, and complete and varied sentences.)
k. may contain a few errors in the correct use of mechanics (i.e., underlining, quotation marks, commas, semicolons, apostrophes, capitalization, and spelling), but errors do not detract from overall delivery

SCORE 3

The student's response

a. addresses the specific writing prompt (partially develops the topic)
b. contains a stated or implied main idea (thesis)
c. shows a sense of audience and purpose
d. contains a minimum of three indented (or clearly delineated) paragraphs
e. has a clear beginning, middle, and end
f. has a main idea developed by supporting details, but these are not consistently well elaborated
g. exhibits some logical order; sequences most steps or ideas with transitions
h. contains appropriate language, but word choice may be repetitive (grade-level vocabulary)
i. maintains a consistent point of view
j. may contain occasional errors in grammar/usage that may detract somewhat from the delivery (Grammar/usage includes subject-verb agreement, verb tense, pronoun case and reference, and complete and varied sentences.)
k. may contain some errors in the correct use of mechanics (i.e., underlining, quotation marks, commas, semicolons, apostrophes, capitalization, and spelling) that may detract somewhat from delivery

SCORE 2

The student's response

a. addresses the specific writing prompt (minimally develops the topic)
b. contains a vaguely implied main idea (thesis)
c. shows little regard for audience and/or purpose
d. may not exhibit indented (or clearly delineated) paragraphing
e. has a beginning, middle, and end
f. addresses the main idea with minimal supporting details
g. exhibits some evidence of organization but does not sequence steps consistently and/or does not use transitions
h. contains some appropriate language; word choice is repetitive (grade-level vocabulary)
i. may not maintain a consistent point of view
j. may contain frequent errors in grammar/usage that may impede communication (Grammar/usage includes subject-verb agreement, verb tense, pronoun case and reference, and complete and varied sentences.)
k. may contain frequent errors in the correct use of mechanics (i.e., underlining, quotation marks, commas, semicolons, apostrophes, capitalization, and spelling) that may impede communication

SCORE 1

The student's response

a. attempts to address the writing prompt (may digress from the topic)
b. does not contain a main idea (thesis) or contains only an implied focus on the topic
c. shows no regard for audience and/or purpose
d. may not exhibit indented (or clearly delineated) paragraphing
e. may lack a beginning, middle, and/or end
f. contains vague or no details that support the topic
g. lacks organization; presentation is rambling and repetitive
h. contains vague and imprecise language (below grade-level vocabulary)
i. does not maintain a consistent point of view
j. exhibits serious errors in grammar/usage that may severely impede communication (Grammar/usage includes subject-verb agreement, verb tense, pronoun case and reference, and complete and varied sentences.)
k. contains repeated errors in the correct use of mechanics (i.e., underlining, quotation marks, commas, semicolons, apostrophes, capitalization, and spelling) that may severely impede communication

SCORE 0

The student's response

a. is incomprehensible
b. is insufficient to score