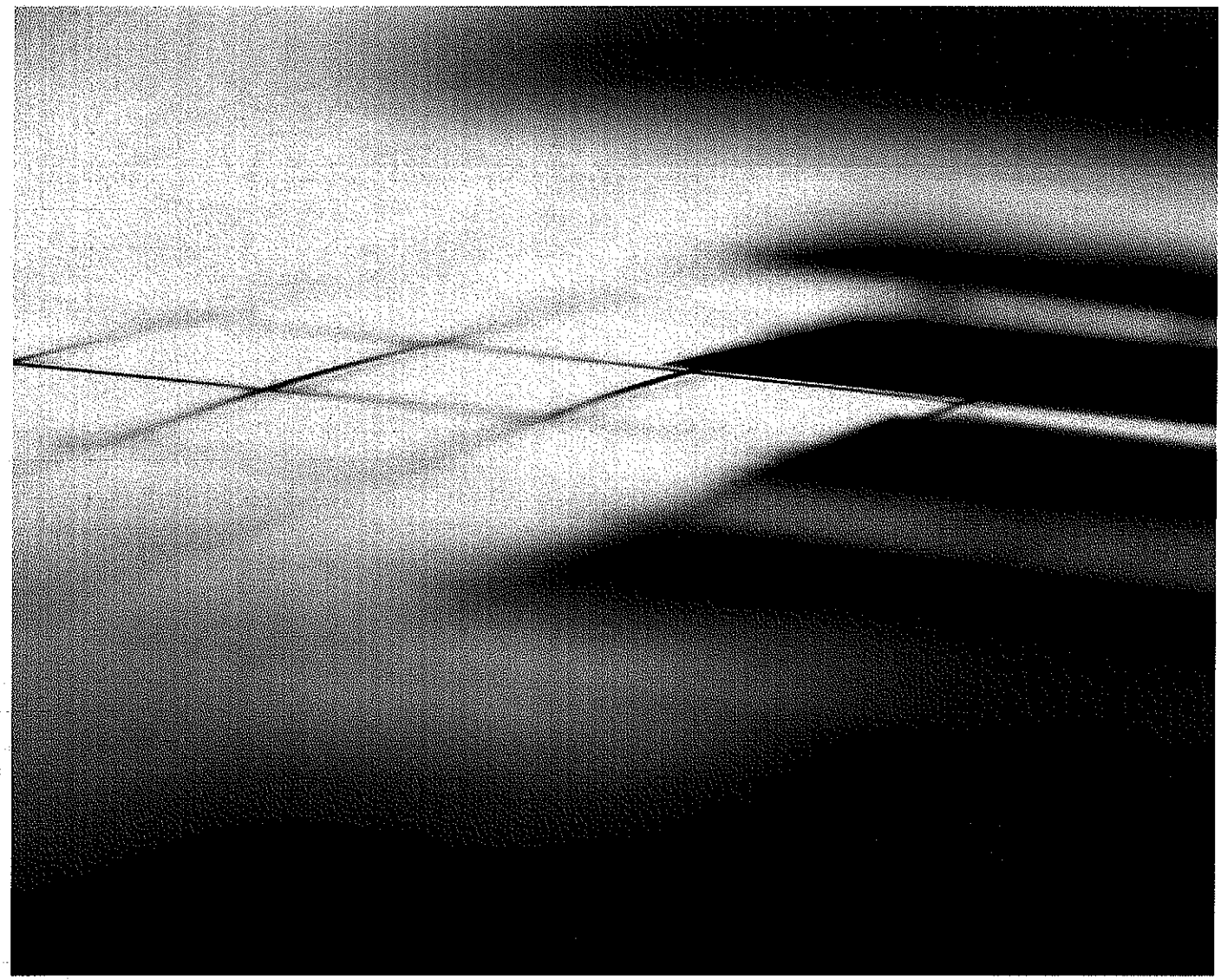
Center for Research in Educational Policy

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

2006-2007 Evaluation of Supplemental Educational Services in Kentucky:

Student Achievement Analysis in Jefferson County Public Schools



Center for Research in Educational Policy

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

2006-2007 Evaluation of Supplemental Educational Services in Kentucky:

Student Achievement Analysis in Jefferson County Public Schools

January 2008

Marco A. Muñoz
Jefferson County Public Schools

Steven M. Ross
Center for Research in Educational Policy

Executive Summary

Purpose of Study

The present study evaluated the effects of SES programs on student achievement in Jefferson County Public Schools (JCPS) in Louisville, Kentucky as part of an overall state-wide SES evaluation for the 2006-07 school year. In an initial study, conducted in JCPS in 2005-06, SES students were compared to matched control students on the Kentucky Core Content Test in Reading and Mathematics, with no overall significant effects associated with an individual SES provider or providers.

Major Findings and Conclusions

The present study replicated the quasi-experimental design (SES students vs. matched control students) employed in 2005-06. The major conclusions supported by the present findings are:

- The overall effects of SES across providers were nonsignificant and small in comparison to similar non-SES students. For Reading, the overall effect size was -.05, indicating virtually no group differences; for Mathematics, the effect size was +.12, indicating a small-to-moderate directional (but nonsignificant) advantage for SES students.
- Regular vs. non-regular attendance at SES sessions had no impact on Reading achievement. However, there was a directionally positive but nonsignificant advantage for students with regular attendence in Mathematics instruction.
- Individual provider results in Reading were mostly inclusive. Of the 17 providers that served 15 or more students, only two were associated with significant effects: Huntington Learning produced positive effects (ES = +.34), and Lincoln Foundation

produced negative effects (ES = -1.48). Of the 17 providers examined, 8 had effects that were positive in direction and 6 that were negative in direction.

- Individual provider results in Mathematics were also mostly inclusive. Of the 8 providers that served 15 or more students, only 1 was associated with significant effects—Sylvan produced positive effects (ES = +.35). Of the 8 Mathematics providers, 3 had effects that were positive in direction and 5 that were negative in direction.

In judging individual providers, it is important to consider that the present quasi-experimental design does not permit control over student selection. Thus, it is entirely possible and very likely that certain providers served students who, due to personal characteristics, family situations, and school experiences, were more at risk than others. Such factors cannot be equalized across providers via statistical adjustments. Further, sample sizes for individual providers were relatively small overall, and subject to error variance associated with multiple grades and schools. As the evaluation is replicated each new year, the aggregated outcomes should yield a more reliable picture of the individual provider effects.

2006-2007 Evaluation of Supplemental Educational Services in Kentucky: Student Achievement Analysis in Jefferson County Public Schools

The No Child Left Behind (NCLB, 2001) Act is sweeping federal legislation designed to help all children become proficient in reading and mathematics by 2013-2014. A core aspect of NCLB is the offering of Supplemental Educational Services (SES). Through SES, students who (1) attend Title I schools in the second year or more of school improvement and (2) participate in the free/reduced-price meal program are eligible to receive this extra academic assistance at no cost to parents (Ross, Potter, & Harmon, 2006). The SES program offers tutoring before or after school hours, on weekends, and during the summer, by state-approved providers, including community-based, faith-based, and private organizations.

Evaluation Purpose

In concert with its emphasis on accountability for schools, NCLB (2001) explicitly requires all states to evaluate the effectiveness of each SES provider in raising student achievement. The present study is part of a larger state-wide evaluation of SES in Kentucky in compliance with federal requirements to examine the effectiveness of SES providers at least once every two years. Because JCPS has the largest number of SES students in the state, has unique characteristics as a large urban district, and has ready access to student participation, achievement, and demographic data; the data for JCPS was analyzed separately. When similar results become available for identical providers from other districts, aggregate effect sizes to indicate state-wide provider effectiveness can be computed. Based on this rationale, an initial study was conducted in JCPS in 2005-06, in which SES students were compared to matched control students on the Kentucky Core Content Test (KCCT) in Reading and Mathematics (Potter, Ross, Munoz, Paek, & McKay, 2007). No significant effects were associated with

individual SES providers or providers overall. The primary purpose of the present study was to extend the analysis of student achievement analyses to the 2006-07 school year, As with the 2005-2006 study, other evaluation goals were to determine the perceptions of varied SES participants and stakeholders—district coordinators, principals/site coordinators, teachers, and parents—towards SES activities, implementation processes, and educational outcomes. This section of the report deals with achievement only.

The major research question regarding student achievement was addressed via a quantitative analysis of student-level scores on state-mandated tests in reading and mathematics. A randomized field trial (RFT), potentially the most rigorous research design, could not be implemented due to the SES program's emphasis on parents' ability to choose whether to enroll their children in SES and which providers to choose, should they enroll. Accordingly, as in 2005-06, the next most rigorous option (Potter et al., 2007), was adopted – a quasi-experimental design (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002) using closely matched program and control students with multiple student-level covariates.

Research Context

The context for the current study was all SES participants in the Jefferson County Public Schools (JCPS) in Louisville, Kentucky. The district is the 26th largest school district in the nation. JCPS is located in a large metropolitan area and has 150 schools serving more than 96,000 students; 33 of these schools were required to offer SES during the 2006-2007 school year. SES is an expensive component of the Title I budget, with cost estimates of approximately $4.5 million for a school year. JCPS educates a high percentage of at-risk urban students with high poverty levels (i.e., over 53% subsidized meals, 54% single-parent homes). The district has a student assignment plan based on managed choice, which facilitates the racial desegregation of

its schools by providing students with transportation from their home neighborhoods to other parts of the district.

The present study is a continuation of the study conducted in 2005-06 of SES implementation in JCPS. Specifically, Muñoz, Potter, and Ross (in press) analyzed student-level achievement scores for the SES and matched control students on the Kentucky state assessment system in Reading and Mathematics [i.e., Kentucky Core Content Test (KCCT)] and school district measures [i.e., demographics, Predictive Assessment Series (PAS) Test]. As previously indicated, findings showed no significant advantages for SES in Reading and Mathematics for any individual provider or for all providers combined. Importantly, with regard to interpreting results from the present study, the Muñoz et al. findings indicated that implementation of SES in the JCPS school district was characterized by (a) limited duration of the tutoring activity relative to regular school programs; (b) lack of sensitivity of high-stakes, standardized tests for assessing gains on specific knowledge skills that may be taught (Linn & Miller, 2005); (c) failure to implement the interventions fully; and (d) absence of effective communication by providers with teachers and parents.

Method

The analysis of program impacts on criterion-referenced test scores was the focus of the study and basis for the methodology described below. Given the relatively small sample sizes for provider *X* grade-level groups, it was decided to replicate the Year 1 (2005-06) procedure of aggregating data across all 25 service providers in an initial analysis. Subsequent analyses examined individual provider effects where sample size was 15 or more students.

Participants and Matching Procedure

District data identified the SES providers from which each eligible student received services during the 2006-2007 school year. Providers serving students in Jefferson County ranged from large national companies to local community-based organizations. A typical tutoring session lasted 1 hour after school, two days per week. Provider programs had a variety of instructional methods including one-on-one or small-group instruction; home-based tutoring, and online instruction. Most programs lasted for several weeks, with the majority of tutoring occurring during the second (spring) semester of the school year.

The SES provider data were combined with databases containing student demographics, formative assessment results, and state assessment results for students in Grades 2-11. Of the 16,684 students who were eligible to receive SES services, 3,381 (20.3%) had valid state assessment data in Reading and/or Mathematics; however, only 2,637 (15.8% of total eligible) had both state assessment and previous test scores. Of these, 1,772 students (10.6%) actually received services, whereas 865 students applied for services but did not participate in SES. The shrinkage in sample size is due largely to the characteristics of the state assessment in Kentucky. Only students in grades 3-8 and 10 take the annual state assessment in Reading, whereas only those in grades 3-8 and 11 are tested in Mathematics. Because no previous achievement baseline assessment scores were available for third graders in both subjects, and eleventh graders in Mathematics, these students were not included in the analyses.

Reading participants. The sample included SES participants and a demographically matched sample of SES-eligible, non-participating "control" students. The participating service providers began implementing SES services in the 2006-2007 school year. Table 1 illustrates the key characteristics by which the treatment and control students were matched at the beginning of

the school year for the analysis of Reading outcomes. As shown, the matches were based on four variables, including previous diagnostic test scores in Reading, gender, race, and participation in the free or reduced-price lunch program. All students included were active participants in the free or reduced-price lunch program and were comparable on grade level. The comparability of the remaining matches was evaluated using chi-square tests for categorical variables [gender, χ (1, N = 2540) = .48, p = .49; race, χ (1, N = 2540) = 9.75, p = .01] and ANOVAs for continuous variables [previous test scores, F (1, 2538) = 9.94, p < .01]. As indicated in Table 1, the comparison sample tended to include more minority students and scored slightly over 2 points lower (d = -.13) than did the SES sample. As a result, we included these variables as covariates in the posttest analysis.

TABLE 1. Treatment and Comparison Students in Reading Matched on Key Characteristics (N = 2,540)

	SES Students				*Comparison Students*			
	M	*SD*	*n*	*%*	*M*	*SD*	*n*	*%*
Previous Test	50.06	17.67	1,697		47.71	17.71	843	
Gender								
Female			840	49.50			405	48.04
Male			857	50.50			438	51.96
Race								
Minority			1086	64.00			592	70.23
Non-Minority			611	36.00			251	29.77
Grade								
3			60	3.50			35	4.20
4			160	9.40			64	7.60
5			175	10.30			60	7.10
6			482	28.40			259	30.70
7			357	21.00			200	23.70
8			359	21.2			158	18.70
10			104	6.10			67	7.90

Note. Only students with complete demographic and testing data were included in the analysis. An aggregated matching procedure was utilized.

Mathematics participants. Table 2 presents the key characteristics by which the treatment and control students were matched at the beginning of the school year (2006-07) for the analysis of Mathematics outcomes. All students included were active participants in the free or reduced-price lunch program and were comparable in grade level. As in the case of the

Reading sample, the comparability of the remaining matches on the three key variables was evaluated using chi-square tests for categorical variables [gender, χ (1, N = 648) = .97, p = .33; race, χ (1, N = 648) = 8.31, p = .01] and ANOVAs for continuous variables [previous test scores, F (1, 646) = 1.97, p = .16]. As found for the Reading sample matching analyses, the comparison sample for Mathematics tended to include more minority students than did the SES sample. Thus, we again included this variable and previous test scores as covariates in the posttest analysis.

TABLE 2. Treatment and Comparison Students in Mathematics Matched on Key Characteristics (N = 648)

	SES Students				*Comparison Students*			
	M	*SD*	*n*	*%*	*M*	*SD*	*n*	*%*
Previous Test	45.13	15.17	468		43.24	15.93	180	
Gender								
Female			249	53.21			88	48.89
Male			219	46.79			92	51.11
Race								
Minority			281	60.04			130	72.22
Non-Minority			187	39.96			50	27.78
Grade								
3			135	28.85			57	31.67
4			159	33.97			64	35.56
5			174	37.18			59	32.78

Note. Only students with complete demographic and testing data were included in the analysis. An aggregated matching procedure was utilized.

Instrumentation

Achievement measures. The achievement measure employed was the KCCT in Reading and Mathematics, with each subtest consisting of 24 multiple choice and 6 constructed response items. The "diagnostic" test, used for matching SES and control students on prior achievement and as the pretest covariate, was the Predictive Assessment Series (PAS), which consisted of 24 multiple-choice items. The PAS ThinkLink benchmark test (ThinkLink, 2007) is considered a reliable predictor of student performance on criterion referenced tests. ThinkLink PAS results can be used to gauge progress toward proficiency levels as defined by each state under the

requirements of NCLB. KCCT Reading and Mathematics contribute 14.5% of the 100% accountability formula used as part of the annual state assessment (Kentucky Department of Education, 2005). All criterion (KCCT Reading and Mathematics) tests were group-administered and scored following standardized procedures.

Only the students who had complete data encompassing the PAS and KCCT Reading and Mathematics served as the basis for the student- level matching component (Rossi, Freeman, & Lipsey, 1999). The PAS was administered at the beginning of the school year to students in grades 3-10 in Reading and to students in grades 3-5 in Mathematics. At the end of the 2006-2007 school year, students in grades 3-8 and 10 took the KCCT Reading subtest, and those in grades 3-8 and 11 took the KCCT Mathematics subtest. Eleventh graders were not included in the present study due to the lack of a baseline test in Mathematics. A summary of the tests and analyses employed is provided in Table 3.

TABLE 3. Summary of Analyses Comparing SES to Control Students on Student Achievement

Outcome Measure	Year	Analysis	Covariate(s)
Grades 3-8, 10 KCCT Reading	2006-2007	ANCOVA	Prior achievement on PAS (beginning of Grades 2-8, 10) and race
Grades 3-5 KCCT Mathematics	2006-2007	ANCOVA	Prior achievement on PAS (beginning of Grades 3-5) and race

Note. KCCT = Kentucky Core Content Test; PAS = Predictive Assessment Series.

Design and Procedure

The research design employed a matched treatment-control pre-posttest design (Cook & Campbell, 1979; Rossi et al., 1999; Shadish et al., 2002). To control for prior achievement, student-level data were analyzed via ANCOVA, with the treatment condition as the between-subjects factor. The PAS pretest scores and race were used as covariates.

Results

Table 4 displays a summary of the KCCT Reading and Mathematics student achievement analyses. The direction of the means in Mathematics slightly favored the SES treatment group over the control group.

KCCT Reading Test Scores

Tests of the homogeneity of regression indicated that assumptions of homogeneity were met and that usage of analysis of covariance (ANCOVA) was justified. The ANCOVA performed on KCCT Reading posttest scores yielded significance for the PAS [$F(1, 2536) = 16.35, p < .01$] and race [$F(1, 2536) = 6.20, p < .05$] covariates. More importantly, the Program effect [$F(1,2536) = 1.69, p = .19$] was nonsignificant, indicating no differences between the SES and the comparison students.

As a follow-up, a sub-group analysis by number of service hours was conducted, using previous test scores and race as covariates. The service hour variable was derived using a median split for the entire sample, whereby the duration of participation by regular attenders > 18.25 hrs. per year and by non-regular attenders < 18.25 hrs. The ANCOVA performed on KCCT Reading by group (i.e., regular attendee, non-regular attendee, comparison group) again yielded significance for the PAS [$F(1, 2535) = 16.32, p < .001$] and race [$F(1, 2535) = 5.86, p < .05$] covariates. More importantly, an overall significant effect was found for treatment groups [$F(2, 2535) = 4.45, p = .01$]. However, post-hoc analyses indicated that the statistically significant difference was between regular ($M_{adj} = 679.05$, SD = 158.03) and non-regular ($M_{adj} = 703.33$, SD = 169.40) attendees, but surprisingly favoring non-regular attendees (Mean Difference = 24.28,p < .05). Comparison students did not differ from either group. The effect

sizes for regular and non-regular attenders relative to the comparison group were -.09 and +.06, respectively.

KCCT Mathematics Test Scores

Tests of homogeneity of regression confirmed that assumptions to justify ANCOVA were met for Mathematics. Accordingly, an ANCOVA using the baseline (Fall 2006) PAS scores and race as covariates was conducted on the KCCT Mathematics subtest scores. Results indicated that the PAS [$F(1, 644) = 29.06, p < .001$] was a significant covariate; race however was not found to be a significant [$F(1, 644) = 0.72, p = .40$] covariate. More importantly, the Program effect was not significant [$F(1, 644) = 1.52, p = .22$], but effect sizes directionally favored the Program group (see Table 4).

As performed for Reading, a sub-group analysis by number of service hours based on the median participation hours (i.e., regular attendee > 23.50) was conducted, using previous test scores and race as covariates. The ANCOVA performed on KCCT Mathematics by group (i.e., regular attendee, non-regular attendee, comparison group) yielded significance for the PAS covariate [$F(1, 643) = 29.09, p < .001$], but not for race [$F(1, 643) = .69, p = .41$]. More importantly, an overall significant effect was not found [$F(2, 643) = 1.09, p = .34$]. However, KCCT mathematics test scores were directionally higher for regular attendees (M = 443.74, SD = 81.24), followed by non-regular attendees (M = 437.23, SD = 81.66), and comparison students (M = 430.25, SD = 82.02). Relative to the comparison group, the effect sizes for the regular and non-regular attendees were +.19 and +.17, respectively.

TABLE 4. Unadjusted and Adjusted Means and Standard Deviations for SES and Matched-Control Comparisons on KCCT Scale Scores in 2006-2007

Comparison Group and Test	*N*	*M*	M_{adj}	*SD*	*ES*	ES_{adj}
2006-2007 KCCT Reading						
SES	1697	685.41	685.18	161.29	-0.05	-0.05
Control	843	693.62	694.07	164.71		
2006-2007 KCCT Mathematics						
SES	468	441.71	440.94	81.34	+0.14	+0.11
Control	180	430.25	432.25	82.02		

Service Provider Analyses

ANCOVA procedures controlling for prior achievement were conducted to assess the impact of specified service providers on Reading and Mathematics test scores. Analyses were only conducted for service providers with a total sample size equal to or greater than $n = 15$ in the target subject. Application of this criterion yielded 17 providers in reading and 8 in mathematics (see Table 5). The analytical procedure included the use of dummy-coded variables for each provider. Most of the variance was captured by the statistically significant previous achievement covariate (PAS), but significant results were obtained in several analyses, as described below.

In *Reading*, 2 of the 17 providers were associated with statistically significant effects:

- Huntington Learning ($n = .256$, $M_{adj} = 710.64$, SD = 149.91) had a *positive* adjusted effect size of +.34, $F(1,509) = 8.07$, $p = .01$).
- Lincoln Foundation had a negative adjusted effect size of -1.48 ($p < .001$).
- Sylvan showed a positive trend (ES = +.18) that approximated significance ($p = .06$).

In *Mathematics*, 1 of the 7 providers with sufficient sample size was associated with statistically significant results:

- Sylvan (n = .256, M_{adj} = 444.14, SD = 81.36) had a *positive* adjusted effect size of +.35, $F(1,189) = 4.51$, $p = .04$).
- Huntington showed a positive trend (ES = +.47) that approximated significance ($p = .07$).

TABLE 5. 2006-07 Service Providers in Reading and Math (n = > 15 students)

Service Provider	Fstats	Fprob	ADJ-M Effect Size
Reading			
Huntington Learning	8.07	.01	+.34
Sylvan	3.69	.06	+.18
A Plus Tutoring	.68	.41	-.18
A to Z In-Home Tutoring	.01	.92	+.02
ATS Educational Consulting	.03	.87	+.10
Bates Community Development	.12	.73	-.19
Brainfuse Online Instruction	.23	.64	-.35
Catapult Online	.13	.72	-.04
Club Z!	.00	.99	-.00
Dr. Lee Epstein	.41	.53	-.43
Learning Disability Association	.16	.69	-.28
Lincoln Foundation	10.09	.00	-1.48
Urban League	.47	.50	+.27
Porter Education	.04	.85	+.07
SOAR	2.71	.10	+.38
The Salvation Army	1.00	.32	-.32
Yes! All Students Can Learn	.07	.80	+.11
Mathematics			
Sylvan	4.51	.04	+.35
Huntington Learning	3.40	.07	+.47
A Plus Tutoring	.12	.73	-.13
A to Z In-Home Tutoring	.27	.61	-.19
Catapult Online	1.64	.20	+.26
Club Z!	.53	.47	-.15
Lincoln Foundation	1.90	.18	-.85
The Salvation Army	.10	.75	-.13

Summary and Conclusions

Similar to other recent evaluations of SES (Potter, Ross, Paek, Pribesh, & Nunnery, 2006; Potter, Ross, Muñoz, Paek, & McKay, 2007; Muñoz, Potter, & Ross, in press), the present results showed only isolated achievement advantages for the SES participants on state-mandated test scores in reading and mathematics. Specifically, on the KCCT Reading subtest in grades 3-8 and 10, and the KCCT Mathematics subtest in grades 3-5, SES participants across the 25 providers serving JCPS were not significantly different than the matched control students. The overall effect sizes were -.05 and +.12, in Reading and Mathematics, respectively, indicating small differences. However, there were isolated significant effects for three individual providers: Huntington Learning was associated with positive impacts and Lincoln Foundation with negative impacts in Reading; Sylvan had positive effects in Mathematics.

One possible explanation for the absence of effects is the limited duration of the tutoring activity relative to regular school programs and other educational experiences. A second, related explanation concerns the lack of sensitivity of high-stakes, standardized tests for assessing higher-order learning or gains in specific knowledge skills that may be taught (Linn & Miller, 2005). But a third, even more direct cause of limited program impacts would be failure to implement the interventions strongly or fully. Although implementation quality was not evaluated directly in the present study, the perceptions of the district coordinator and school principals/SES coordinators did not show strong support for provider services at schools within the district.

The present findings are qualified by several factors. First, due to the inability to employ a randomized experimental design, firm causal conclusions cannot be reached. Sampling representativeness is limited by student choice in program enrollment and participation. As a

result, it is difficult to isolate the impact of the SES program participation from other possible explanations for student achievement changes. For example, the participation and outcomes of students might be affected by teachers or parental involvement. Despite confounding variables, the systematic matching procedure and statistical controls should have established sufficiently strong internal validity for attributing differential group change more to the interventions than to potentially confounding intervening factors (Cook & Campbell, 1979; Shadish et al., 2002).

A second limiting factor was the shrinkage of sample sizes for both the SES and control groups due to incomplete testing data. Direct year-to-year comparisons of KCCT scale scores or proficiency levels are difficult because the reading and mathematic tests are given to particular grade levels. Our results indicate that only a small subset of SES and comparable students had test scores on the accountability grades. Differences of this sort could limit to whom we might generalize our results and, thus, might compromise the study's external validity.

In summary, the major conclusions supported by the present findings are:

- The overall effects of SES across providers were nonsignificant and small in comparison to similar non-SES students. For Reading, the overall effect size was -.05, indicating virtually no group differences; for Mathematics, the overall effect size was +.12, indicating a small-to-moderate directional (but nonsignificant) advantage for SES students.
- Regular vs. non-regular attendance at SES sessions did not improve Reading achievement. However, there was a directional but nonsignificant advantage for regular attenders in Mathematics.
- Individual provider results in Reading were mostly inclusive. Of the 17 providers that served 15 or more students, only two were associated with significant effects:

Huntington Learning produced positive effects (ES = +.34), and Lincoln Foundation produced negative effects (ES = -1.48). Of the 17 providers in the analysis, 8 had effects that were positive in direction and 6 that were negative in direction.

- Individual provider results in Mathematics were also mostly inclusive. Of the 8 providers that served 15 or more students, only 1 was associated with significant effects—Sylvan produced positive effects (ES = +.35). Of the 8 Mathematics providers, 3 had effects that were positive in direction and 5 that were negative in direction.

In judging individual providers, it is again important to consider that the present quasi-experimental design does not permit control over student selection. Thus, it is entirely possible and very likely that certain providers served students who, due to personal characteristics, family conditions, and school experiences, were more at risk than others. Such factors cannot be equalized across providers via statistical adjustments. Further, sample sizes for individual providers were relatively small overall, and subject to error variance associated with multiple grades and schools. As the evaluation is replicated each new year, the aggregated outcomes should yield a more reliable picture of the individual provider effects.

References

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Kentucky Department of Education. (2005). *CATS interpretative guide.* Frankfort, KY: Kentucky Department of Education.

Linn, R. L., & Miller, M. D. (2005). Measurement and assessment in teaching. Upper Saddle River, NJ: Pearson Education.

Muñoz, M. A., Potter, A. P., & Ross, S. M. (In Press). Supplemental Educational Services as a consequence of the NCLB legislation: Evaluating its impact on student achievement in a large urban district. *Journal of Education for Students Placed At Risk.*

No Child Left Behind Act. (2001) Pub. L. No. 107-110.

Potter, A., Ross, S. M., Muñoz, M. A., Paek, J., & McKay, D. (2007). *Supplemental Educational Services in the State of Kentucky: 2005-2006.* Center for Research in Educational Policy, University of Memphis, Memphis, TN.

Potter, A., Ross, S. M., Paek, J., Pribesh, S. L., & Nunnery, J. A. (2006). *Supplemental Educational Services in the State of Louisiana: 2004-2005.* Center for Research in Educational Policy, University of Memphis, Memphis, TN.

Ross, S. M., Potter, A., & Harmon, J. (2006). *Evaluating Supplemental Educational Service providers: Suggested strategies for states.* Memphis, TN: Center for Research in Educational Policy, American Institutes for Research, and Academic Development Institute.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Newbury Park, CA: Sage.

Shadish,W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin Co.

ThinkLink. (2007). *Research*. Retrieved March 27, 2007 from http://www.thinklinklearning.com/research.php.