Mid-Western Educational Researcher

Volume 21 | Issue 4 Article 5

2008

The Effect of Supplemental Educational Services on Student Learning Outcomes

Jane Beese
The University of Akron

The Effect of Supplemental Educational Services on Student Learning Outcomes

Jane Beese
University of Akron

Abstract

This study investigated the effects of Supplemental Educational Services on student achievement in reading and math. A meta-analysis was conducted to investigate the relationship between student achievement gains in reading and participation in SES. Six studies were used in the meta-analysis with a total of 395 participants. The combined weighted effect size correlation was r = 0.48. A meta-analysis was conducted to investigate the student achievement gains in math. Two studies were used in the meta-analysis, with a total of 136 participants. Their combined weighted effect size correlation was r = 0.04. The magnitude of this effect size correlation is trivial, indicating that participation in math SES produces little effect; however, the effect size correlation associated with reading SES suggests a large effect on learning outcomes in the content area of reading.

Introduction

In the 20th century, educational reform played a central role in the war on poverty. The Elementary and Secondary Education Act of 1965 (ESEA), a Federal law largely defined by Title I, focused on lessening the disparities in academic performance between poor and wealthy schools by promoting equitable academic achievement (Stullich, McCrary, & Roney, 2006). Initially federal funds were used by states to equalize school funding and ensure alike treatment but as time went by with little change and growing concern about international competition, this focus evolved into a standards-based accountability system with the objective of equal outcomes for all students (Bowe, Cronin, Kingsbury, & McCall, 2005).

Two presidential education initiatives have preceded the current efforts of NCLB. The first was Goals 2000: Educate America Act of 1994. Goals 2000 established a framework in which to identify world-class academic standards, to measure student progress, and to provide the support that students may need to meet the standards (Carr, 2001). By 2000, the goals were for all children starting school to be ready to learn, to increase high school graduation rates to 90%, for students in grades 4, 8, and 12, to demonstrate competency in specified subject areas, for all students to be ready to enter the workforce, for students to rank first in the world in math and science, for every school to be free of drugs, and for teachers to have access to professional development. The National Education Goals Report demonstrated modest improvements in several goals, which included more children born with a healthier start in life, more families reading and telling stories to their children, improved math scores for students in fourth and eighth grades, and more degrees earned in math and science. In other areas, the nation regressed; 12th grade reading achievement declined, fewer secondary school teachers held degrees in their subject areas, the gap in college completion rates between White and Hispanic students widened, and school violence increased. Other areas showed no change.

In 1994, the US legislature passed a second presidential initiative, the School-to-Work Opportunities Act (STW). This law required states to coordinate school-to-work plans with the educational reforms they were already planning with Goals 2000. Both acts involved a restructuring, rescheduling, and rethinking of educational practices—in other words, a systemic change within education. The failure of STW can be linked to the deficiency in understanding the process of implementing curriculum change or change in traditional subject areas. Suggested improvements were to develop a concentrated effort to influence pre-service teacher education programs within the higher education structure, a definitive body of knowledge, a scope and sequence for the new material, professional development models, an administration involvement plan, and an evaluation plan using subjective normative testing. According to Carr (2001), "Future federal educational efforts should be able to improve from the STW shortfalls and create a more effective design and implementation methodology" (p. 34). It has been suggested that educational reform should be developed from programs that provide information and statistics so that knowledgeable decisions can be made (Kruse, forthcoming).

No Child Left Behind

Historically, the No Child Left Behind Act of 2002 (NCLB) is the largest kindergarten through 12th-grade federal education program. NCLB represents a restructuring and redirection of federal efforts to support elementary and secondary education. Prior to NCLB, public school accountability had been a state and local responsibility with the federal government and national organizations playing a supportive role. In the history of education reform, there is no federal law that exceeds the nationalization of education policy such as NCLB (Elmore, 2004). The federal government has become highly involved in the daily operation of public education by instituting a federal law that imposes a single accountability system determined suitable for all schools while setting national parameters on state and local

accountability systems (Bowe, Cronin, Kingsbury, & McCall, 2005; Cocoran & Goertz, 2005; Elmore, 2004). As a result of the expanding federal role, NCLB has raised fundamental issues regarding who controls education (Sunderman, Kim, & Orfield, 2005).

Behind the actions taken on behalf of NCLB lie four principles. First, the law provides an accountability system that identifies underperforming schools. A fundamental component of accountability is testing; NCLB put in place a system of testing to validate school effectiveness and holds schools and districts accountable for student learning. Many educators advocate this type of approach; i.e., one that designates a specific and well-planned curriculum and uses tests to determine the extent the curriculum has been taught and learned (Glickman, Gordon, & Ross-Gordon, 2007; Ravitch, 1985; Walberg, 2001).

Second, the law increases the opportunities for students who are in schools that are deficient in the subject areas of reading and math. Two school options given to parents in underperforming schools are school choice and Supplemental Educational Services (SES). Through these options, students are provided with more opportunity to get the help they need.

Third, the law increases the capacity for students to become proficient. NCLB policymakers have set a target rate of 100% proficient by 2014. By improving the overall effectiveness of the public school system, the path is potentially cleared for all students to become proficient.

Fourth, the law attempts to reduce the achievement gap among subgroups of the population. A key feature of NCLB is the goal of narrowing the achievement gap between white and minority students by using a federally-required set of measures. States and school districts must report on the progress of specific subgroups of students (e.g., racial/ethnic, impoverished, disabled, and limited-English proficient).

Under NCLB, every state is required to set standards of academic content and measure each student's yearly progress in the core subjects of reading, math and science. Adequate yearly progress (AYP) is the determinant of school improvement. A school that fails to meet AYP is classified as in "improvement" status. Schools identified for improvement are subject to a series of sanctions. Sanctions are based on markets and privatization theories (Sunderman, Kim, & Orfield, 2005). Schools in the first year of improvement are required to offer to all students the option to transfer to another school not in improvement status. Students in schools in the second year of improvement are eligible for Supplemental Educational Services.

Both public school choice and Supplemental Educational Services are provisions of NCLB intended as corrective action for schools identified as needing improvement. Parents of eligible students are notified by their child's school and may select from a list of qualified providers (Kruse, Liang, & Beese, 2005, 2006; Kruse, Liang, & Widenbaugher, 2004). Students who come from low-income families attending Title I schools, whether or not they performed at proficiency, are eligible for SES (Corwin & Wilhelm, 2006). In circumstances where only a limited number of students can be provided services due to financial constraints or other limitations, priority is given to the lowest achieving students.

Supplemental Educational Services

Supplemental Educational Services, otherwise known as SES, are educational activities provided outside the normal school day, designed to enhance the educational services that are provided during a regularly scheduled school day (Kim & Sunderman, 2004; Smole, 2006a). In simple terms, Supplemental Educational Services (SES) are after-school tutoring programs. Since 2003, the SES school option has afforded students in failing schools access to tutors where they can get assistance in reading and math. SES providers can vary by type. Providers can be district or school programs or private tutoring businesses. All providers must be approved by each state's Department of Education and aligned with the state's reading and math content standards.

The money to finance SES is provided by redirecting Title I funding. The former goal for Title I was to provide educational opportunities for identified groups of disadvantaged children. This goal has now shifted to individualized access to educational programs. NCLB requires districts to set aside 20% of their yearly Title I allocated funds for school choice transportation and SES (Kim & Sunderman, 2005). Districts in improvement status are required to spend a minimum of 5% of the total set aside on SES. According to the U.S. Department of Education (2005e), "The per-child cost for Supplemental Educational Services is determined by dividing the district's Title I, Part A allocation by the number of children residing within the district aged 5 to 17 who are from families living below the poverty level." Costs paid out for SES are strictly related to provider fees. Title I funds are not permitted to be used for administrative costs or any other cost associated with implementing SES within a school or district. This controversial aspect of SES funding has resulted in claims that NCLB is an unfunded mandate.

Many educational leaders and state and local officials are critical of NCLB policies. In an effort to establish accountability, schools have drastically narrowed their curriculum, becoming intently focused on teaching reading and math (Fletcher, 2005). The system seems unresponsive to problems and views all the schools the same. Furthermore, fiscal considerations may discourage districts from promoting NCLB's choice options. The more that students pursue these options, the more districts will have to devote the mandatory 20% Title I budget set aside to SES programs rather than to programs already in existence with the likelihood that even 20% will not be adequate to cover the cost (Finn & Hess, 2004).

SES was established to be regulated by the state and local school districts in an effort to increase individual students' academic achievement through after-school tutoring for students in schools classified as needing improvement. As with any new program, monitoring and evaluation, data

collection and analysis are vital to the overall success of SES. Fortunately for SES providers, forerunners in SES implementation have identified "best practices" (Cohen, 2003; U.S. Department of Education, n.d.e).

Research Studies

Very few studies have investigated the effects of NCLB on student learning outcomes. Due to the paucity of research in this area, each study shall be discussed in some detail. A study conducted by Bowe, Cain, Kingsbury, and McCall (2005) used the Growth Research Database from the Northwest Evaluation Association to compare student achievement and student growth on a common and reliable scale. The participants included hundreds of thousands of students in school districts across the country. The purpose of the study was to provide an initial view of the law and to identify trends. Findings from the studies indicated that state level tests tend to improve observed achievement and there is evidence that NCLB has improved student achievement since its adoption, although the effect is smaller than the testing effect. The measured growth in achievement may not necessarily be due to interventions on behalf of NCLB but may be attributed to the process of testing and retesting students and/or regular academic growth.

While NCLB has shown positive effects on student achievement and growth, there are two concerns raised by this study. The first is that at the current rate of change, schools will not be close to reaching the requirement of 100% proficiency by 2014. The second is that students in ethnic groups that have demonstrated achievement gaps in the past have had less growth under NCLB, and demonstrate less growth in comparison to European-American students with the same baseline score. NCLB was in its initial stages and it may have been too early in program implementation to identify the extent to which NCLB will influence educational change in the future.

Secretary Spellings (U.S. Department of Education, n.d.b) asserts NCLB has had the intended positive effect on students. The latest Nation's Report Card (U. S. Department of Education, 2006a) shows steady growth and gains by students particularly among younger and minority students; overall fourth grade and eighth grade math scores increased as well as fourth grade reading scores. African-American and Hispanic fourth graders reached the highest reading and math scores for their groups than in any previous year, and African-Americans and Hispanic eighth graders reached the highest math scores for their groups than in any previous year. In both fourth and eighth grades, a higher percentage of white, African-American, Hispanic and Asian/Pacific Islander students performed at or above proficient than those in previous years. Although improvement has been measured, the results suggest middle and high school students may be an area of weakness.

While NCLB has garnered vast interest and many articles, there has been very little written about the SES portion of the initiative. Basically there are two areas of study pertaining to SES. The first area investigates SES implementation and the second area looks at student achievement gains. One study presented key findings over a period of three years (David et al., 2006). Data were collected from a variety of sources including: a yearly survey of 1,300 district Title I administrators, a yearly survey of 739 principals, yearly case study visits, interviews of Title I administrators, and analyses of state accountability system components. From the study emerged five themes:

1. Small district schools were more likely to exit improvement status than large districts.
2. Participation in school choice remained at one percent and participation in Supplemental Educational Services increased from 7 to 19%.
3. An increased number of states provided technical assistance to schools in improvement status.
4. Strategies for school improvement remained similar across the three years nationally.
5. School poverty and district size were higher predictors of exiting improvement status than improvement strategies.

The biggest challenges districts who implemented SES faced included the lack of available providers (especially in small, rural districts), communication with parents, and assessing provider performance. In both 2002-03 and 2003-04, the number of eligible and participating students was substantially increased in urban and very large districts from 9,000 to 16,000 (on average) students. SES providers were primarily non-faith-based and non-online providers.

A study by Kim and Sunderman (2004) used 11 urban districts from a geographically, politically, and demographically diverse sample to provide a wide range of local contexts in which to examine the ability of districts to implement SES. The results of the study confirmed that SES was not widely used during the first year. The demand for services was low, primarily due to the inconvenience of services being offered outside of regular school hours and away from eligible students' neighborhoods. The first year also documented tremendous administrative burdens faced by districts with no increased funding. Moreover, there was growing concern of the potential for SES to fragment Title I, seriously disrupting other school reform efforts by diverting resources away from the neediest students.

Another study (Anderson & Laguarda, 2005) used case studies conducted during 2003-04 school year and followed baseline data that were collected in the previous year. The study also conducted interviews in a purposive sample of six states and nine school districts, which were selected because they appeared to be relatively advanced in the process. Findings indicate that after two years, states, districts, schools, and providers were overcoming some of the initial trials of SES implementation. A noted area of improvement was establishing routines for reviewing applications and getting a list of providers out to districts sooner. District administrators continued to confront additional administra-

tive responsibilities and were in the process of developing systems that would streamline operations. Other challenges included moderate increases in participation, evaluating provider performance, improving communication with parents, managing administrative costs, and payment to providers when student attendance is irregular.

In a study conducted by the Chicago Public School Office of Research, Evaluation and Accountability (2005), baseline achievement levels of students who participated in the program were compared to eligible students who did not participate. The students with tutoring increased at national norms, while those students without tutoring had slightly fewer students at national norms. In addition to measuring student gains, gains between providers were examined. Students from one specific provider were shown to outperform students from other providers. Few researchers have examined the impact of Supplemental Educational Services on student achievement and school performance. There is no body of research that provides conclusive evidence documenting the effect of SES on learning outcomes for low-income and minority students. Does student participation in SES increase student learning outcomes in reading and math? The earliest implementation of SES was during the school year 2003-2004 and the impacts of which are in their infancy.

Few studies have been published that report the effects of Supplemental Educational Services on student outcomes, primarily because traditional statistical methods typically used in evaluations just won't work well with SES data. The purpose of this study is to implement an evaluation approach based on contemporary statistics that can effectively analyze SES data, a meta-analytic approach using effect size analysis. Our effective implementation of this evaluation approach may provide important information for future program evaluations and previously administered SES programs. Specifically, this study will add to the body of knowledge on educational reform to better understand the effectiveness of various SES programs. With any policy change, it is important to monitor the impact of that change and to evaluate the value of associated new programs. The primary focus of this study was to investigate the relationship between SES and achievement gains in a typical sample of SES students within Ohio and to identify strengths and weaknesses in provider practices that may be associated with student learning gains.

Methodology: Analyses and Results

To determine if student achievement is enhanced by student participation in SES, the researcher examined three years of results from an ongoing statewide Evaluation Project and performed a meta-analytic analysis that estimated the overall impact of Supplemental Educational Services in predicting increased student achievement. As with all SES programs, the participants in the study were selected on the basis of need, and achievement performance data were collected at the start and at the end of SES programs.

The data groups for this study come from two public school districts located in the State of Ohio. SES programs used in this study were implemented by local school district personnel. Due to the eligibility requirements for SES participation, it is assumed that individual SES samples are generally similar to SES populations nationwide, primarily comprised of high poverty, low achieving, minority students (U.S. Department of Education, 2005e). SES samples have been determined to be similar to SES populations nationwide (U.S. Department of Education, 2005e). Sufficient research and conclusive statistical evidence indicating SES populations are similar demographically in composition (U.S. Department of Education, 2005e). Students are eligible, whether or not they perform below the established level of proficiency, poorly on an assessment, or are part of a particular subgroup who attend Title I schools that have not met AYP for two consecutive years (Corwin & Wilhelm, 2006). National studies have characterized SES students as coming from low-income families, high poverty schools, and within the lower rankings for statewide assessments (U.S. Department of Education, 2004b, 2005e). Based on a theoretical framework and current research, it is assumed that the SES providers are generally comparable to other SES providers across the state and across the country.

The data were in the form of learning gains on academic assessments in reading and math. Individuals with incomplete records such as a pretest score and no posttest score were eliminated. A *t* test for dependent means was conducted to determine the statistical significance of the difference between the pre- and posttests for each group, see Table 1 for results. The data were analyzed with both descriptive and inferential statistics to determine irregularities in distribution. Means, standard deviation, skewness, and kurtosis were investigated and reported for critical variables.

Researchers are challenged regarding the use of statistical power and analysis when sample sizes are varied. According to Cohen (1992), "It is most useful to determine the *N* necessary to have a specified power for given (significance criterion) and ES" (p. 156). This problem frequently arises in ex post facto studies like the ones used in this study. In order to detect differences, as the sample sizes decreased in size the level of risk became greater.

An effect size functions as a descriptor to explicate the meaning of effect size. Cohen (1988, p. 25) described effect size as, "the average percentile standing of the average treated (or experimental) participant relative to the average untreated (or control) participant. At statisticians' disposal are two types of effect sizes; standardized difference and variance-accounted for effect sizes (Thompson, 2000). While an effect size correlation is calculated using original standard deviations, the Cohen's d uses a pooled standard deviation to calculate effect. There is reason to consider the impact of selecting effect size type. Dunlop, Cortina, Vaslow, & Burke (1996), Thompson (2000), and Rosenthal (1993) reported that pooled standard deviations tend to inflate the actual effect. Therefore, for this investigation, original standard deviations were used to compute effect size correlations.

In ex post facto research, it is necessary to determine the effect size (ES) in order to set the Type I error. The ES

was used to quantify the effectiveness of SES because it clarified the results and explained how well SES treatment worked (Cohen, 1988; Rosnow & Rosenthal, 1996). The ES allows results to be interpreted beyond statistical significance to practical implication and determine whether or not the results add to the general body of knowledge (Cohen, 1992). Statistical significance alone does not reveal the true size of the effect. The effect of SES on student achievement in math and reading was presumed to be large or noticeable to the observer.

In order to minimize the risk of Type I error, the investigator examined the necessary sample sizes suggested by Cohen (1992) for the specified level of significance and effect sizes. The suggested sample size for *t* test of dependent means with a large ES and level of significance .10 was 20 participants. The Ns for the groups with the level of significance of .10 were 11, 6, 46, and 17. From 2002 to 2004, SES participation increased from 7% to 19% (David et al., 2006). The rate of increase was reflected in this study's sample size. Groups F and H met the required sample sizes. In spite of insufficient sample sizes of groups A, B, D, E, and G, it was decided to proceed with the analysis. There were six reading groups (A through F) and two math groups (G and H), see Figure 1. The effect size (ES) for a large number of participants is assumed to be a more precise estimate of the population of the effect size based on a much smaller population. Because the sample sizes of the groups were so discrepant, each group was weighted by its number of participants.

Descriptive statistics and *t*-tests results for dependent mean comparisons, defined by pre and posttest mean differ-

Table 1
Pre- and Post-Test Distribution

Sample	Subject	*N*	Measure	Mean	Std. dev.	Skewness	Kurtosis
A	Reading	11	Pretest	54.09	13.99	–.78	–.23
			Posttest	47.64	15.92	.96	.14
B		6	Pretest	353.33	7.00	.50	1.74*
			Posttest	382.00	30.35	–.06	1.74*
C		46	Pretest	225.43	77.04	–.25	–.42
			Posttest	271.96	75.14	.09	–.41
D		8	Pretest	4.75	1.03	–.39	–.49
			Posttest	5.50	1.06	.47	–.83
E		12	Pretest	11.83	8.10	.91	–.06
			Posttest	18.75	11.87	–.04	–1.46*
F		312	Pretest	2.66	1.69	.50	–.05
			Posttest	3.31	1.78	.91	–.55
G	Math	17	Pretest	7.47	6.85	2.48*	7.20*
			Posttest	12.00	9.51	1.45*	2.09*
H		119	Pretest	3.52	1.59	–.52	–.56
			Posttest	3.49	1.33	–.52	.15

Note: *Indicates skewness or kurtosis outside of the acceptable 1.2 range.

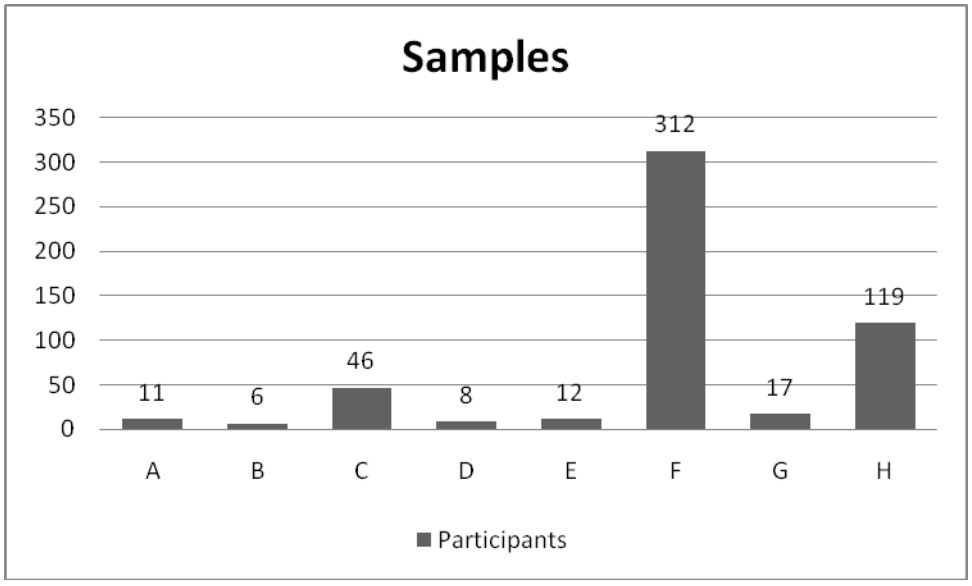

Figure 1. Sample distribution.

ences for each group, are reported in Table 2. Groups A and H do not show significant pre-post differences. Groups B, C, D, E, F, and G show significant differences, $p < .10$. The results of these analyses indicate significantly ($p < .10$) greater achievement post scores on average for the SES groups.

In order to conduct an empirical evaluation and aggregate statistics, it is necessary to convert all of the summary statistics of the various studies into a common effect size. Using the statistics provided from the *t* tests of dependent means, *t* values were converted into Pearson Product Moment Correlations *r*. Where the original result indicated a negative effect, the posttest had a larger mean than the pretest; the *r* was treated as negative.

Confidence intervals for each sample's weighted effect size were computed and comparisons made to determine the stability of their effect size. Confidence intervals for every sample included zero, suggesting the distinct possibility of no effects. Because individual study sample sizes were relatively small, the level of significance was set at .05. Table 3 provides a summary index of the combined procedures of statistical significance of the *t*-test results.

Next, the effect size correlations were combined by averaging the raw Pearson correlation coefficients. The combined effect size correlation (*r*) for reading = .71 and math = .01. At this point in the meta-analytic process, typically each *r* is transformed into a *Z* statistic using Fisher's *r* to *Z* transformation and then the *Z* scores are averaged and transformed back into *r* values. There is some controversy involving this process. Fisher (1932) argued that *Z* scores somewhat overestimate *r* when sample sizes are small and *r* is large. A more conservative method (Hunter et al, 1982 and Rosenthal, 1984 & 1993) is to use a weighted average *r*. A weighted *r* is calculated by weighting each correlation by the number of subjects in that particular study. Taking into consideration Fisher's (1932) argument, the weighted *r* was used to combine effect sizes. The results of the combined weighted effect sizes (r^1) were reading = .48 and math = –.04, results are shown in Table 3.

The heterogeneity of ES was examined to discern whether or not it was appropriate to synthesize the study results into one meta-analysis or if subsets should be considered. The most contributing factor in heterogeneity was sample size. To find the heterogeneity of the studies, the researcher calculated Q statistics and then distributed the scores in a chi-square. A chi-square was used to determine the degree of probability that the observed variance in ESs was the result of sample error alone. The criterion was set at .05. For reading the chi square was significant ($\chi^2(5) = 12.67$, $p < .05$). The critical value for chi-square with $df = 5$ and $p < .05$ is 11.07. The χ^2 is greater than the critical value indicating the variability across effect sizes does exceed what would be expected by just chance alone. This may be due to the degree of variation between sample sizes amongst studies. Sample F (with 312 subjects) contributes the most

Table 2
Paired Samples t Tests for Dependent Means

Sample	*N*	Pre-test X	*sd*	Post-test X	*sd*	*t*	Dir. of effect	*p*	ES
A	11	54.0	13.9	47.6	15.9	1.17	–	.26	.21
B	6	353.3	7.0	382.0	30.3	2.27	+	.07*	.94
C	46	225.4	77.0	271.9	75.1	5.40	+	.00*	.39
D	8	4.7	1.0	5.5	1.0	4.58	+	.00*	.37
E	12	11.8	8.1	18.7	11.8	3.40	+	.00*	.32
F	312	2.6	1.6	3.3	1.7	19.73	+	.00*	.21
G	7	7.4	6.8	12.0	9.5	3.02	+	.00*	.27
H	119	3.5	1.5	3.4	1.3	0.81	–	.41	.04

Note. *statistical significance at .10.

Table 3
Standard Values for the Samples

Sample	Subject	*N*	*df*	*t*	ES *r*	wES *r'*	Lower Upper Confidence intervals
A	Reading	11	10	1.17	–0.35	–0.01	–0.69 to 0.69
B		6	5	2.27	0.71	0.01	–1.13 to 1.14
C		46	45	5.40	0.63	0.07	–0.29 to 0.31
D		8	7	4.58	0.87	0.02	–0.86 to 0.00
E		12	11	3.40	0.72	0.02	–0.65 to 0.67
F		312	311	19.73	0.75	0.59	–0.08 to 0.15
G	Math	17	16	3.02	0.60	0.08	–0.46 to 0.49
H		119	118	0.81	–0.07	–0.06	–0.19 to 0.18

to the heterogeneity of the studies. The studies are considerably heterogeneous, making it quite likely that a Type I error influenced the end result for the reading analysis. Due to the uneven sample sizes, a Type I error may indicate differences when in actuality there are none. A chi-square was calculated to determine the heterogeneity of math samples. The critical value for chi-square with $df = 1$ and $p < .05$ is 3.84. The χ^2 value for math is 0.15 and does not exceed the critical value indicating the distribution of effect sizes is homogenous and neither of the two groups needs to be removed.

Confidence intervals for ES and weighted-ES were computed and comparisons made to determine the stability of the effect size. The level of significance was set at .05. The confidence intervals for reading were 0.3791 to 0.5809 and do not include zero. Therefore, the probability that there is likely a true effect in reading exists. The confidence intervals for math were –0.2119 to 0.1319 and do include zero. Thus, it is possible that the math effects may be a fluke, unlikely to replicate.

One shortcoming of a meta-analysis is that a researcher only has access to studies that have been published. Because an unknown amount of research remains unpublished and/or may end up locked in a file cabinet somewhere, this is known as the file drawer problem (Borg, Gall, & Gall, 2003). In order to combat this issue regarding the stability of results, one must calculate the Fail-safe N. The researcher computed the number of nonsignificant studies that would have to be hidden away or filed away in order to make the meta-analysis nonsignificant, see Table 4. If the Fail-safe N is larger than the determined number of studies, then the meta-analysis is considered valid. The Fail-safe N for reading was 812 studies and the critical number of nonsignificant studies 40. There would have to exist at least 812 nonsignificant studies to render the meta-analysis for reading insignificant because that number exceeds the critical number therefore no file drawer problem exists. The number of significant studies that would need to exist for math to be significant was 2 and the critical number of nonsignificant studies 20. The critical number exceeds the Fail-safe N and therefore the study suffers from a file drawer problem. It must be noted that none of the samples in this study have been published and in effect this study is the Fail-safe N. Table 4 provides a summary of the meta-analyses.

Discussion

This study yielded relevant and applicable findings regarding student participation in SES, student achievement. A meta-analysis was conducted to investigate the relationship between student achievement gains in reading and participation in SES. Six reading studies were used in the meta-analysis, with a total of 395 participants. A χ^2 was conducted to determine the heterogeneity of effect sizes. The results were $\chi^2(5) = 12.67$, $p < .05$ which is statistically significant. The samples were determined to be heterogeneous, most likely due to the difference in sample sizes. Sample F contributed most to the heterogeneity of the study. Although precautions were taken and the samples were weighted, the uneven distribution of participants and samples may contribute to rejecting the null when there is no true effect, a Type I error. The combined weighted effect size correlation = 0.48 and is considered a large effect. It is improbable the results are due to chance. It would have required an additional 812 nonsignificant studies to render the meta-analysis nonsignificant. The file drawer problem does not affect this study because only 40 additional nonsignificant studies are likely to exist. The CI range –0.38 to 0.58 includes zero within its interval. Therefore, there is the possibility that the effect size is equal to zero, suggesting the possibility of no effects.

This study yielded relevant and applicable findings regarding student participation in SES, student achievement, and provider practices. A meta-analysis was conducted to investigate the student achievement gains in math experienced by students who participated in SES. Two studies were used in the meta-analysis, with a total of 136 participants. A χ^2 was conducted to determine the heterogeneity of effect sizes. The results were $\chi^2(1) = 0.15$ that indicated the studies were homogeneous. The combined weighted effect size was $r = 0.04$, which Cohen (1992) considers insignificant. This may be in effect a Type II error and due to the paucity of studies to compare in the meta-analysis.

Participants in SES are predominantly from low-income families, high poverty schools, rank in the lower percentile on statewide assessments, and many are minority students (U.S. Department of Education, 2005). The students that are eligible for SES are a homogenous group most likely because our schools are failing these types of students. Schools have an obligation to teach all students and SES may be more suitable for this student demographic. The successes of SES programs should be examined and viable elements should be introduced into the regular classroom to assist this student demographic.

Educational practice is plagued with attempts to develop specific approaches to teaching that assume the approach will be effective in any setting and for all types of students

Table 4
Meta Analysis Summary Table

Variable	No.	*N*	*r*	r[1]	wCI Lower Bound	wCI Upper Bound	Fail-safe N	Critical No.
Reading	6	395	.71	.48	0.3791 to 0.5809		812	40
Math	2	136	.01	.04	–0.2119 to 0.1319		2	20

(Danielson, 1996). Investigating the variables associated with SES may provide data regarding the elements associated with academic gains. For example, if a lower student to teacher ratio in math is found to be a predictor of math learning gains, a school district may choose to incorporate breakout math classes for its students who are falling behind. In many districts, gifted children are sent to breakout classrooms for math or reading. These classes often have smaller teacher to student ratios. Offering the same type of program for remedial students may have the same effects as SES with the benefits of funding remaining within the school and the students' "school day" not being extended.

In conducting this study, certain observations have been made that could serve as the catalyst for future research. Data do not distinguish between regular academic growth and SES treatment. As it stands now, the numbers reflect regular academic growth with SES treatment. Without more accurate measures, we can only suggest what the effectiveness of NCLB policies might be, but we cannot be certain. Gains may be attributed to other factors such as new technology, professional development, or a new math series. Without proper data collection efforts in all states, the proficiency and effectiveness of the SES sanction remains undetermined.

Another area of focus for future research may be the differential effect between math and reading. Historically, there has been more extensive research done in the area of reading than in math. This contributes to the widespread and accepted research-based learning strategies and instructional practices in reading. This may be a contributing factor in the effectiveness of SES programming.

Are children better off having received SES? That question needs further analysis in order to justify the cost to public schools and the redirecting of Title I funding. Since the program is relatively new and there have been reported difficulties in implementation during the first 2 years, the researcher suggests further investigation regarding the impact of SES on both mathematics and reading scores in a longitudinal study. The study should examine learning gains in both reading and mathematics in an effort to determine the point of most dramatic affect on student achievement as well as long-term gain.

References

Anderson, L. M., & Laguarda, K. G. (2005). Case studies of supplemental services under the no child left behind act findings from 2003-2004. *U.S. Department of Education*. 1-85.

Borg, W., Gall, J., & Gall, M. (2003). *Educational research: An introduction* (7th ed.). Boston: Pearson Education.

Bowe, B., Cronin, J., Kingsbury, G., & McCall, M. (2005). *The impact of the NCLB act on student achievement and growth: 2005 edition.* Northwest Evaluation Association.

Carr, F. (2001). A critique of school-to-work (stw) integration into traditional education subjects. *The Journal of Economics & Economic Education Research*. Arkansas State University: Whitney Press.

Chicago Public Schools Office of Research, Evaluation & Accountability (2005). SES tutoring programs: An evaluation of the second year-part one of a two-part report. *Department of Evaluation and Data Analysis*.

Clark-Carter, D. (1997). *Doing quantitative psychological research from design to report.* East Sussex, UK: Psychology Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cohen, J. H. (2003). Supplemental education: Six essential components. *Principal, 82*(5), 34-37.

Corcoran, T., & Goertz, M. (2005). The governance of public education. In S. Fuhrman & M. Lazerson (Eds.), *The public schools* (pp. 25-56). New York: Oxford University Press.

Corwin, T., & Wilhelm, S. (2006). *Supplemental educational services web cast*. Retrieved June 25, 2006, from U.S. Department of Education.

David, J., Laguarda, K., Lash, A., Lopez-Torkos, A., Padila, C., Shields, P., Skolnik, H., & Woodworth, K. (2006). *Title I accountability and school improvement from 2001 to 2004* (pp. 1-160). Washington, DC: U.S. Department of Education.

Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*, 170-177.

Elmore, R. (2004). *School reform from the inside out, policy, practice, and performance.* Cambridge, MA: Harvard Education Press.

Finn, C., Jr., & Hess, F. (2004). Inflating the life rafts of NCLB: Making public school choice and supplemental services work for students in troubled schools. *Phi Delta Kappan*, *86*(1), 34-58.

Fisher, R. A. (1932*). Statistical methods for research workers* (4th ed.). London: Oliver and Boyd.

Fletcher, M. A. (2005). Federal mandate: Is the No Child Left Behind act making the grade? *Crisis, 112*(5), 16-19.

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2007). *Supervision and instructional leadership: A developmental approach.* (7th ed.). Pearson Education.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Kim, J., Orfield, G., Sunderman, G. L., & Tracey, C. A. (2004). Listening to teachers, classroom realities and no child left behind. *The Civil Rights Project by Harvard University.*

Kim, J., & Sunderman, G. L. (2004). Increasing bureaucracy or increasing opportunities? School district experience with supplement educational services. *The Civil Rights Project by Harvard University.*

Kim, J., & Sunderman, G. L. (2005). Measuring academic proficiency under the no child left behind act: Implications for educational equity. *Educational Researcher, 34*(8), 3-13.

Kruse, S. (forthcoming). *The educational leader as decision maker: Processes, issues and considerations.*

Kruse, S., Liang, X., & Beese, J. (2005). Evaluation report of supplemental education services. *External Evaluation Report to Ohio Department of Education*.

Kruse, S., Liang, X., & Beese, J. (2006). Evaluation report of supplemental education services. *External Evaluation Report to Ohio Department of Education*.

Kruse, S., Liang, X., & Widenbaugher, B. (2004). *Supplemental evaluation services external evaluation project report.* Akron, OH: The University of Akron.

Ravitch, D. (1985). *The schools we deserve: Reflections on the educational crises of our times*. New York: Basic.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Rosenthal, R. (1993). *Meta-analytic procedures for social research*. London: Sage.

SES: The Basics. Retrieved June 25, 2006, from www.tutorforkids.org./basics.asp

SES Pilot Programs. Retrieved June 25, 2006, from www.ed.gov

SES-Who Does What? Retrieved June 25, 2006, www.tutorsforkids.org

Smole, D. P. (2006a). Supplemental educational services for children from low-income families under ESEA title I-A. *CRS Report for Congress*. Congressional Research Service: The Library of Congress.

Stullich, S., Eisner, E., McCrary, J., & Roney, C. (2006). National assessment of Title I: Interim report. Volume I: Implementation of Title I. *Institute of Education Sciences*, Washington, DC: National Center for Educational Evaluation and Regional Assistance.

Sunderman, G. L., & Kim, J. (2004). Increasing bureaucracy or increasing opportunities? School district experience with supplemental educational services. *The Civil Rights Project Harvard University*.

Sunderman, G. L., Kim, J., & Orfield, O. (2005). *NCLB meets school realities: Lessons from the field*. Thousand Oaks, CA: Corwin Press.

Thompson, B. (2000). *A suggested revision to the forthcoming 5th edition of the APA publication manual*. Retrieved August 30, 2008, from http://www.coe.tamu.edu/~bthompson/apaeffec.htm

U.S. Department of Education (2002). *No Child Left Behind: Supplemental educational services, non-regulatory guidance.* Washington, DC: Author.

U.S. Department of Education (2003). *No Child Left Behind*. Washington, DC: ED Pubs, Publications Center.

U.S. Department of Education (2004a). *Supplemental educational services policy.* Washington, DC: ED Pubs, Publications Center.

U.S. Department of Education (2004b). *Creating strong supplemental educational services programs* (pp. 1-63). Washington, DC: ED Pubs, Publications Center..

U.S. Department of Education (2005a). *Commission Staff Research Report, Growth models: An examination within the context of NCLB. Commission on No Child Left Behind*. The Aspen Institute. Retrieved September 5, 2006, from www.nclbcommission.org

U.S. Department of Education (2005b). *Nation's report card shows continued progress*. Retrieved June 27, 2006, from www.ed.gov.

U.S. Department of Education (2005c). *No Child Left Behind: What parents need to know.* Washington, DC: Department of Education.

U.S. Department of Education (2005d). *No Child Left Behind: A roadmap for state implementation.* Washington, DC: Department of Education,.

U.S. Department of Education (2005e). *No Child Left Behind: Supplemental educational services non-regulatory guidance.* Washington, DC: Department of Education.

U.S. Department of Education (2006a). *The facts about… Measuring progress*. Retrieved 2006 from www.ed.gov/print/nclb/accuntability/ayp/testing.html

U.S. Department of Education (2006b). *Digest of education statistics.* Washington, DC: National Center for Educational Statistics.

U.S. Department of Education (n.d.a). Budget proposal advances NCLB, introduces competitiveness initiative. *The Achiever*, *5*(3).

U.S. Department of Education (n.d.b). *Key policy letters signed by the education secretary or deputy secretary*. Retrieved June 25, 2006, from www.ed.gov/print/elsec/guid/ secletter/060515.html

U.S. Department of Education (n.d.c). *Letter to chief state school officers, supplemental educational services policy: LEA conditions on providers*. Retrieved October 2006, from http:///www.ed.gov/policy/elsec/guid/stateletters/choice/ses082604.htlml

U.S. Department of Education (n.d.d). *Office of elementary and secondary education and office of innovation and improvement, supplemental educational services (SES) policy*. Retrieved May, 2006, from http://www.ed.gov/policy/elsec/guid/stateletters/ choice/ses051006.html

U.S. Department of Education (n.d.e). *Policy and program studies service report highlights: Early implementation of supplemental educational services under the No Child Left Behind act: year one report*. Retrieved 2006 from www.ed.gov/rschstat/eval/disadv/ supplementalyear1

Walberg, H. J. (2001). Achievement in American schools. In T. Moe (Ed.), *A primer on American schools* (pp. 43-98). Stanford, CA: Hoover Institution Press.

Wilhelm, S., Rees, N., & Corwin, T. (2006). Supplemental educational services web cast.