

Implementation and Outcomes of Supplemental Educational Services: The Tennessee State-Wide Evaluation Study

Steven M. Ross
Allison Potter
Jangmi Paek
Dawn McKay

*Center for Research on Educational Policy
The University of Memphis*

William Sanders
James Ashton
SAS

Supplemental Educational Services (SES), a component of the No Child Left Behind (NCLB) Act, provides free tutoring to economically disadvantaged children who are attending Title I schools in their 2nd or more years of school improvement. This research evaluated SES in Tennessee to determine the: (a) impacts on student achievement, and (b) perceptions of SES implementation and outcomes by teachers, district coordinators, principals/site coordinators, and parents. Using value-added methodology, statistical analyses of achievement data controlled for both student ability and teacher effects in 2 alternative models. Not surprisingly, parent reactions to SES were highly positive, whereas those by the 3 other stakeholder groups were more mixed. Achievement results from both analytical models yielded mostly small and nonsignificant provider effects. The implications of the findings for evaluating SES are discussed with regard to both research and policy issues. Recommendations are offered for broadening the evaluation of SES through smaller mixed-methods studies to examine implementation and educational outcomes in more highly controlled contexts.

Supplemental Educational Services (SES) are a core provision of the Elementary and Secondary Education Act, as reauthorized by the No Child Left Behind Act (NCLB, 2001). Through SES, students from low-income families are eligible to receive extra academic assistance if they attend a Title I school that has not made Adequate Yearly Progress (AYP) for at least 3 years (i.e., is in its second or more years of school improvement). The services primarily or exclusively take the form of individual or small-group tutoring in reading and/or mathematics, which must take place outside of regular school hours. Each state is responsible for approving SES providers to offer services in reading and math. Approved providers indicate the districts in which they are interested in working, and the districts, in turn, assume responsibility for disseminating the listing of available providers to the families of eligible students. Although there have been concerns about low participation rates, which have been estimated to include only 20% of SES-eligible students nationally (Center on Education Policy, 2006; Sunderman, 2006), it is noteworthy, nonetheless, that close to 250,000 students were served in 2005–2006 (Ascher, 2006; U.S. Department of Education, 2006), with an estimated increase up to 500,000 in 2006–2007.

Given the extensive scope, implementation demands, and costs of SES, the need for rigorous and comprehensive evaluation is essential (Burch, Steinberg, & Donovan, 2007; Ross, Potter, & Harmon, 2006; Sunderman, 2006). In this regard, an additional responsibility of states, as stipulated by NCLB (2001) policies, is to evaluate provider effectiveness after 2 years and remove from the approved list those who fail to (a) increase students' achievement for 2 consecutive years, or (b) provide services consistent with applicable federal, state, and local health, safety, and civil rights requirements.

Determining the impact of 30 or 40 hr of tutoring on a child's academic performance is highly challenging. Add to the mix a lack of funding for evaluation, personnel, and expertise, and it is little wonder that most states have been slow to monitor and evaluate provider effectiveness in accord with the NCLB (2001) requirement (Ascher, 2006; Burch et al., 2007). Notably, Tennessee has emerged as a national leader in the SES monitoring and evaluation arena, having been one of only a handful of states to initiate a comprehensive mixed-methods evaluation starting as early as 2003–2004 (Potter, Ross, & McDonald, 2005e) and continuing in 2004–2005 (Potter, Ross, McDonald, Nunnery, & Paek, 2006). In these studies, multiple linear regression (MLR) analyses, controlling for prior-year achievement and selected student demographics (e.g., free or reduced-price lunch, gender, ethnicity), were conducted to compare SES-tutored and nontutored students enrolled in the districts being served. Across both years, there were 42 separate analyses of provider effectiveness in reading and 37 in mathematics. None showed statistically significant effects. Although MLR is one of the recommended and most frequently used approaches in state SES evaluations (see Ross et al., 2006, and discussion in sections to follow), inability to control for teacher and school effects, both highly

influential sources of variance in achievement (Sanders & Horn, 1995; Teddlie & Reynolds, 2000), would weaken the power of MLR to detect SES impacts.

The focus of this research is the 2005–2006 evaluation of SES in Tennessee. As in the two prior studies, the study was designed to examine: (a) impacts on student achievement by the individual providers serving students in the state, and (b) perceptions of SES implementation and outcomes by teachers, district coordinators, principals/site coordinators, and parents. Using value-added methodology (Sanders & Horn, 1995), we were able to increase the rigor of the achievement analyses conducted in the prior years and in other SES studies (as reviewed in the following) by controlling for the effects on achievement scores of both student ability (prior achievement) and teacher effectiveness.

As reviewed by Muñoz, Potter, and Ross (this issue), individual tutoring has long been regarded by researchers and practitioners as one of the most effective ways of adapting instruction to individual differences in school settings (Good & Brophy, 1987; Slavin, 2006; Tingley, 2001; Wasik, 1997). The theoretical and practical rationales are quite straightforward. In virtually all schools, individual teachers routinely face the difficult challenge of trying to accommodate, in a single class, 20–25 students functioning at many different skill levels. Class size reductions to 13–17 students, as demonstrated in the classic Tennessee Student–Teacher Achievement Ratio research (Finn & Achilles, 1999; Nye, Hedges, & Konstantopoulos, 2000), can yield meaningful benefits for student achievement, largely by increasing opportunities for teachers to work individually and adaptively with students. Tutoring, although typically time limited compared to regular classroom instruction, greatly extends the potential for instructional adaptation by reducing the student–teacher ratio to as low as 1:1.

Prior studies of one-on-one tutoring frequently show positive effects on achievement in restricted and controlled contexts consisting of a small number of classrooms or schools (e.g., Fuchs, Fuchs, Yazdian, & Powell, 2002; Green, Alderman, & Liechty, 2004; Ismail & Alexander, 2005; Lauer et al., 2003; Wasik, 1997). Evaluating the effectiveness of statewide implementations of a federally mandated program, such as SES, is much more challenging. The tutoring being evaluated is not confined across providers to a standard instructional orientation delivered by tutors having similar backgrounds or training. The extensive scope of the evaluation context, encompassing multiple providers, districts, schools, and stakeholder groups, greatly reduces the ability of the researcher to control, measure, and adjust for implementation variables (Ross et al., 2006).

Realistically, intermittent tutoring has much less potential to yield measurable effects on student achievement than would more complex and time-intensive interventions such as comprehensive school reform models (Correnti & Rowan, 2007; Desimone, 2002; Kidron & Darwin, 2007) and regular classroom programs in reading and math (e.g., see Slavin & Fashola, 1998). For example, Borman, Hewes, Overman, and Brown (2003) concluded from a meta-analysis of compre-

hensive school reform that the overall effects of the 29 most widely used models were between $d = .09$ and $d = .15$, whereas a synthesis by Borman and D'Agostino (1996) of the results from all federal evaluations of Title I programs yielded an overall effect size of only $d = .11$. Is it reasonable to expect that 30–40 hr of tutoring (comparable in duration to about 6 extra days of school) would produce effects even that high? In the case of SES evaluations, the potential impacts of treatment are likely to be attenuated further from the contamination produced by multiple extraneous variables (e.g., regular school programs and curricula, teacher effects, other after-school programs) and, where state achievement tests are used as the outcome measure, their lack of sensitivity for measuring incremental or specific types of learning growth (Linn & Miller, 2005; Schaps, 2007). It is not surprising that prior evaluation studies of SES (see review by Muñoz et al., this issue) have obtained mostly nonsignificant and small provider effects on student achievement. The most positive findings come from a recent national study (Zimmer, Gill, Razquin, Booker, & Lockwood, 2007) in which participation in SES across seven school districts had a small but significant positive effect ($d = .08$ to $.09$) on students' achievement in reading and math. Students participating for multiple years experienced larger gains. Interpretation of these effects as a benchmark for expected SES impacts should be made cautiously, however, given that (a) the comparison students in the study, although analytically equated via the multiple regression models used, appeared to represent a higher performing population than SES participants; (b) no qualitative or descriptive data were collected on the nature or quality of tutoring interventions; and (c) analyses aggregated data from students serviced by multiple providers.

Given the previously described considerations, the rationale for this study and its methodology is based on several assumptions. First, it remains an open and critically important question, both scientifically and educationally, as to whether SES programs, in individual states and nationally, are positively impacting student achievement. Even if the true effect sizes of SES are small, they should be measurable over time as sample sizes and statistical power increase (e.g., see Zimmer et al., 2007). Therefore, each new study adds to the database for the particular provider and context (e.g., state or school district) and the potential for making valid conclusions about provider effectiveness. Second, because state-mandated achievement assessments are presently the basis under NCLB (2001) for evaluating schools and requiring them to offer SES, it is logical to examine what, if any, impacts of the tutoring experiences are demonstrated on such measures. However, the absence of conclusive evidence would not necessarily indicate that the tutoring experiences are ineffective in engendering educational benefits for other types of learning or psychosocial development. Third, because the measurement of SES effects is challenging for the reasons previously described, it is important to explore analytical methods that potentially provide increased power and precision. Fourth, use of a mixed-method design (Chatterji, 2005; Johnson & Onwuegbuzie, 2004)

can provide insights into how SES was implemented by providers and a basis for interpreting student achievement outcomes.

Clearly, the most salient and potentially influential component of SES evaluations is the design employed to analyze provider effectiveness in raising student achievement. Because this study uniquely analyzed and compared such effects using three alternative designs, it seems relevant to review common methodologies. Ross et al. (2006) recently described, for state SES evaluations, several broad analysis categories, each subsuming more specific variations. At the simplest level, states may choose to perform a *Benchmark Comparison*, which is essentially computation by provider of the aggregate percentages of students attaining proficiency on the state assessment relative to established benchmarks set by the state (e.g., 80% for Reading) in the given year or to a comparison group of nontutored students. This approach was recently employed to initiate examination of provider effects (not to make summative evaluation judgments) in the state of Virginia (Potter et al., 2007). Although applications of the Benchmark Comparison design typically do not analyze student-level data (only aggregate percentages), even if they did, the bluntness of proficiency levels as the dependent measure significantly limit the ability to detect change relative to using interval data such as raw scores or Normal Curve Equivalent (NCE) scores.

A second, much more rigorous category is the MLR design, using pretest and posttest scores obtained from samples of SES participants and nonparticipants from the same or similar Title I schools. Specifically, the MLR model is used to estimate the current year's student-level achievement score based on students' SES participation, prior achievement, and other characteristics (e.g., free or reduced-price lunch status, gender, ethnicity) that are predictive of achievement. A *fixed effect* (effectiveness score) for each provider, and for providers overall, can be estimated and subjected to an inferential statistical test to determine the significance of its difference from zero. This type of approach has been used in the SES evaluations conducted in Louisiana (Potter, Ross, Paek, Pribesh, & Nunnery, 2006), Los Angeles Unified School District (Rickles & Barnhart, 2007; Rickles & White, 2006), Pittsburgh (Zimmer, Christina, Hamilton, & Prine, 2006), and, most recently, Chicago Public Schools (Office of Research, Evaluation, & Accountability, 2007) and the nine-district national study (Zimmer et al., 2007). Although more demanding than the Benchmark Comparison, the MLR approach is still relatively economical with regard to time and cost. Importantly, it increases power and precision for determining provider effects by controlling statistically for the influences of individual student differences, particularly in prior achievement. One disadvantage is a lack of control over teacher and school effects where SES students served by particular providers and comparison students are concentrated in different learning environments. Teacher effects, in particular, tend to account for substantially larger portions of variance in student achievement than do program effects (Ross, Stringfield, Sanders, & Wright, 2003), thus potentially attenuating the ability to detect SES influences.

A third, and generally preferred, category is the matched samples design. The most straightforward approach involves matching each SES student with a similar student attending the same school, according to prior achievement, eligibility for SES, and possibly other individual characteristics such as ethnicity and gender. If a suitable within-school comparison student is not available, the matching search would be extended to comparable Title I neighborhood schools, preferably ones in the second year of improvement status. The important advantage of this approach, especially where there are within-school matches, is controlling for school effects on student achievement. The disadvantage, compared to other designs, is that it is substantially more time consuming to conduct.

In this present study, the SAS Educational Value-Added Assessment System (EVAAS) methodology and databases, which constitute core components of the Tennessee Comprehensive Assessment Program (TCAP), were used for analyzing provider effects in a highly rigorous manner. Specifically, linkages to teacher effects data permitted this major source of variance in student achievement to be controlled statistically by the analytical model. Further, to provide a reliability substudy, our research design also included a matched-pair analysis. Thus, research questions addressed (a) the effects of provider services on student achievement, (b) the consistency of achievement outcomes using alternative analytical models, and (c) perceptions of district and school SES coordinators, parents, and teachers regarding the benefits and implementation of the tutoring services.

METHOD

Overall Design

The design consisted of a descriptive study of SES provider perceptions by involved stakeholders and a quasi-experimental design in which the achievement scores of SES students were compared to those of similar nontutored (or control) students. As will be described in the following, three statistical analysis models were employed: (a) one taking into account students' prior achievement but failing to control for teacher effects as in typical MLR designs employed in prior SES studies (e.g., Office of Research, Evaluation, & Accountability, 2007; Potter et al., 2005e; Potter, Ross, McDonald, et al., 2006; Rickles & White, 2006); (b) a second model nesting teachers within grades as a random effect; and (c) a third model matching similar SES and control students within the same classroom (teacher). The basic design for the descriptive study consisted of surveying the following groups of respondents: (a) the district SES coordinator, (b) principals/site coordinators (to be called *school leader* in this text) in participating SES schools, (c) teachers of students receiving SES, and (d) parents of students receiving SES.

Achievement Study

Achievement measure and sampling. The TCAP is a criterion-referenced multiple-choice test administered each spring. The test is mandated by the State of Tennessee for assessing AYP in the subjects of Reading/Language Arts (R/LA) and Math, in Grades 3-8, in compliance with NCLB (2001) policies. For this achievement analyses, respective end-of-year 2006 TCAP scale scores in each subject (measured in state units) were modeled as a function of a student's predicted score based on 2 years of previous test scores, the student's grade level, the service provider, and the teacher (Sanders et al., 2006). Because this model used predicted scores based on prior TCAP achievement (first available in Grade 3), only data from Grades 4-8 could be modeled. There also were data for some SES students in Grades 9-12, but the sample sizes by provider were much too small to support valid inferences.

The model also restricted the analysis to students (a) having linkages to SAS EVAAS teacher data (Sanders et al., 2006), (b) having predicted scores based on having sufficient prior-year scores, (c) receiving tutoring from providers that tutored at least 20 students in the given subject statewide, and (d) receiving tutoring from one provider only. The requirement of a minimum of 20 students was selected by agreement with the state SES administrators and the researchers to ensure inclusion of a sufficient number of students in the analysis for providing a reliable estimate for a provider. Due to missing data, however, the final sample sizes were sometimes smaller.

Analysis models. For each district, the set of schools that had SES students tutored in 2005-2006 was identified. Students are tested each year in four subjects: Math, R/LA, Science, and Social Studies. Therefore, most students had either 4 or 8 historical scores from the prior 2 years for use as predictors for the target Math and R/LA achievement analyses. The tutoring information was then merged with the student data and each student was matched with 2005-2006 teacher of record.

For modeling purposes, a predicted score for each student was derived from SAS EVAAS methodology (Sanders et al., 2006). The prediction parameters were developed from the sample of students at the subset of schools in each district. The predicted score was, therefore, what the student would be expected to score on the spring Math or R/LA tests, assuming that the student had the average schooling experience, where "average schooling experience" was based on the subset of district schools that had students tutored. The outcome measure used was NCE scores. Provider effects are represented as "estimated" differences between the tutored (SES) and nontutored comparison students in NCE units.

Three models were fitted to these data. As previously indicated, data for only providers having at least 20 tutored students were included for analyses within each model. A second criterion was that following any attrition for missing data,

each provider subsample must contain at least 10 students having the required scores and teacher linkages. Otherwise, the given provider was considered to have insufficient data for analysis. The three models consisted of: the following:

1. Model 1: Fixed effects were predicted score, provider, and grade. There was no teacher effect included in the model, so that the teacher effects and the tutoring effects are confounded. This type of analysis resembles the MLR design that has been frequently used in recent state and district evaluations due to its balancing suitable rigor with high practicality (Ross et al., 2006).

$$\text{Score} = \text{predicted score} + \text{grade} + \text{provider}$$

2. Model 2: Fixed effects were predicted score, grade, and provider. Teachers nested within grade was included in the model as random effect. This model extends conventional SES analyses by accounting for teacher effects on student achievement, thereby estimating provider effects free of teacher influences.

$$\text{Score} = \text{predicted score} + \text{grade} + \text{provider} + \text{teacher}(\text{grade})$$

3. Model 3: Fixed effect was grade level. Block was nested within grade as a random factor. Each block consisted of a pair of matched students, one SES and one comparison. The matching for each SES student was done within the classroom (teacher) by selecting the closest nontutored student based on predicted score differences not exceeding 10 NCE units. Note that it was not possible to find a match for each tutored student due to a variety of factors—missing teacher, clustering of tutored students with no nontutored classmates, or simply not finding a suitable match (when predicted score differences exceeded 10 units). As for Model 2, this model estimates the provider effect free of the teacher effect.

$$\text{Score} = \text{grade} + \text{provider} + \text{block}(\text{grade})$$

Implementation and Participant Satisfaction Analyses

The questionnaires used in this study were adapted from those originally developed by the first and second authors (Ross et al., 2006) and field-tested and validated in the SES evaluations conducted in Tennessee (Potter et al., 2005e; Potter, Ross, McDonald, et al., 2006), and Louisiana (Potter et al., 2006). Because the instruments are nearly identical to those used in Kentucky and described in this issue by Muñoz et al., we will summarize them only briefly below.

The State of Tennessee Supplemental Educational Services District Coordinator Questionnaire. This questionnaire was used to assess the perceptions of SES provider services by the district SES coordinators or local education agencies (LEAs) (Potter, Ross, & McDonald, 2005a). The questionnaire contains 13 Likert-type items evaluating providers in areas such as communication, adherence to federal and local laws, ability to serve both special education and English language learner students, and overall satisfaction. An open-ended comment

section is also included. Most items on the LEA questionnaire are also addressed in surveys for other groups (teachers, principals, and parents), to allow for triangulation.

The State of Tennessee, Supplemental Educational Services Teacher Questionnaire. Similar to the LEA, this instrument contains 11 Likert-type items and an open-ended comment section dealing with provider communications, coordination of services with classroom instruction, and other key aspects of tutoring activities and their associated benefits for students (Potter, Ross, & McDonald, 2005d).

The State of Tennessee, Supplemental Educational Services Principal/Site Coordinator Questionnaire. This questionnaire was used to assess the perceptions of SES by the school leader most familiar with the SES services provided at the school (Potter, Ross, & McDonald, 2005c). It contains 11 Likert-type items and an open-ended section identical to the teacher questionnaire, with wording changes to reflect the different respondent group targeted.

The State of Tennessee, Supplemental Educational Services Parent Questionnaire. This instrument was the only paper-based survey in the evaluation. It contains 7 Likert-type items on topics such as provider communication, progress reports sent home on student goals achieved, and overall provider satisfaction (Potter, Ross, & McDonald, 2005b). It also contains an open-ended section and 3 Likert-type questions about parent satisfaction with district support and communications regarding the SES program.

Procedure

District coordinators were e-mailed login information to take their surveys online during the Spring semester of 2006. Teachers and principals received their online login information in a letter enclosed with the box of parent surveys sent to each school. Parent surveys were sent home with students, completed by parents, and returned to the school for mailing back to the researchers. As described by Muñoz et al. (this issue), return rates for the respondent groups (except for the district coordinators) were difficult to measure, due to the unknown factors of the number of students tutored or how many teachers had SES students in their classrooms when the surveys were distributed. Therefore, the school leader, teacher, and parent responses could be biased in either a positive or negative direction, and thus were interpreted cautiously where not clearly triangulated by other data.

RESULTS

Student Participation by District

Table 1 presents a summary of the number of students in Grades 4–8 who received tutoring in Math and R/LA in the three Tennessee districts offering SES in 2005–2006 (Davidson County, Hamilton County, and Memphis). As shown in the table, (a) eight different providers, six of which tutored Math and seven R/LA, offered SES in Davidson County; (b) eight tutored in both subjects in Hamilton County; and (c) 13 tutored in both subjects in Memphis. The number of students included in the database as receiving tutoring in Math and R/LA, respectively, were 142 and 74 in Davidson County; 378 in Hamilton County; and 383 and 348 in Memphis.

The data used in the analyses are presented in Table 2. As described in the Method section, a provider must have tutored in total (across districts) at least 20 students in a given subject statewide and, of these, at least 10 students had to have both predicted scores and teacher linkages. Following application of this criterion, the number of providers tutoring either or both subjects was 2 in Davidson County, 4 in Hamilton County, and 6 in Memphis.

A summary of the matched pairs available for the Model 3 analyses is provided by district and provider in Table 3. As reflected in the far-right column, successful

TABLE 1
Tennessee Tutoring Evaluation 2005–2006: Summary of Students Receiving Tutoring by District and Provider

| County/Provider | Subject | Grade (N) | | | | | Total (N) |
|---|----------------------|-----------|----|----|----|---|-----------|
| | | 4 | 5 | 6 | 7 | 8 | |
| Davidson County | | | | | | | |
| A's and B's Tutorial Service and Learning Institute | Math | | 7 | 3 | 5 | 4 | 19 |
| A to Z In-Home Tutoring | Math | 6 | 11 | 13 | 6 | 5 | 41 |
| | Reading/ Language | 6 | 11 | 13 | 6 | 5 | 41 |
| Club Z! In-Home Tutoring | Math | | 2 | 1 | | | 3 |
| | Reading/ Language | | 1 | 4 | 2 | 1 | 8 |
| | Math | 21 | 15 | 18 | 10 | 6 | 70 |
| Education Station | Reading/ Language | 6 | | | | | 6 |
| | Reading/ Language | 2 | | | | | 2 |
| Fannie Battle Day Home for Children, Inc. | Reading/ Language | | | | | | |

(continued)

TABLE 1 (Continued)

| County/Provider | Subject | Grade (N) | | | | | Total (N) |
|---|----------------------|-----------|----|----|---|----|-----------|
| | | 4 | 5 | 6 | 7 | 8 | |
| Huntington Learning Centers, Inc. | Math | | 1 | 1 | 1 | 3 | 6 |
| | Reading/ Language | 1 | 2 | 2 | 2 | | 7 |
| Martha O'Bryan Center for Reading Achievement Program | Reading/ Language | 6 | | | | | 6 |
| | Math | 2 | | 1 | | | 3 |
| SkyLearn TN | Reading/ Language | 3 | | 1 | | | 4 |
| | Math | 20 | 13 | 6 | 3 | 3 | 45 |
| Hamilton County A to Z In-Home Tutoring | Reading/ Language | 20 | 13 | 6 | 3 | 3 | 45 |
| | Math | 1 | | | | | 1 |
| ATS Educational Consulting Services—Project Success | Reading/ Language | 1 | | | | | 1 |
| | Math | 3 | 1 | 4 | 2 | 1 | 11 |
| Bethlehem Literacy Academy—After School | Reading/ Language | 3 | 1 | 4 | 2 | 1 | 11 |
| | Math | | 2 | 3 | | 1 | 6 |
| Club Z! In-Home Tutoring | Reading/ Language | | 2 | 3 | | 1 | 6 |
| | Math | | 2 | 3 | | 1 | 6 |
| Educational Enterprises | Math | 1 | | | | | 1 |
| | Reading/ Language | 1 | | | | | 1 |
| HONORS Learning Center | Math | 6 | 5 | 7 | 5 | 4 | 27 |
| | Reading/ Language | 6 | 5 | 7 | 5 | 4 | 27 |
| Newton Learning, a Division of Edison Schools | Math | 8 | 19 | 56 | 9 | 13 | 105 |
| | Reading/ Language | 8 | 19 | 56 | 9 | 13 | 105 |
| The Urban League Street Academy | Math | 83 | 99 | | | | 182 |
| | Reading/ Language | 83 | 99 | | | | 182 |

(continued)

| | | | | | | | |
|---|----------------------|----|----|----|----|----|-----|
| Memphis A's and B's Tutorial Service and Learning Institute | Math | 1 | 2 | 8 | 7 | 18 | |
| | Reading/ Language | | 1 | 2 | 2 | 5 | |
| Back on Track Tutoring Association | Math | 9 | 6 | | | 15 | |
| | Reading/ Language | 8 | 7 | | | 15 | |
| Club Z! In-Home Tutoring | Math | 1 | 2 | 3 | 11 | 5 | 22 |
| | Reading/ Language | 3 | 5 | 2 | 13 | 6 | 29 |
| | Math | 7 | 7 | 2 | | 16 | |
| | Reading/ Language | 9 | 10 | 2 | | 21 | |
| Exceptional Parent Connection | Math | 1 | 1 | | | 1 | 3 |
| | Reading/ Language | 1 | | 1 | | 2 | 4 |
| Failure Free Reading Instant Achievement Center | Math | 6 | 1 | | | 7 | |
| | Reading/ Language | 5 | 5 | | | 10 | |
| Gain and Sustain Learning | Math | | | | 2 | 6 | 8 |
| | Reading/ Language | | 1 | | 4 | 10 | 15 |
| If I Had a Hammer, Inc. | Math | 1 | 2 | 30 | 41 | 56 | 130 |
| | Reading/ Language | 7 | 2 | 20 | 32 | 26 | 87 |
| Innovative Ideas | Math | | | 31 | 49 | 27 | 107 |
| | Reading/ Language | | | 28 | 44 | 33 | 105 |
| Memphis Lemoyne-Owen College Scholars Academy | Math | | 2 | 1 | 3 | 6 | |
| | Reading/ Language | | 2 | 2 | 3 | 7 | |
| PLATO Learning, Inc. | Math | 9 | 5 | 11 | | 25 | |
| | Reading/ Language | 10 | 5 | 10 | | 25 | |
| Success Educational Services | Math | | | 3 | 4 | 7 | |
| | Reading/ Language | | | 2 | 2 | 4 | |
| X-CEL Operating Foundation | Math | | | 2 | 17 | 19 | |
| | Reading/ Language | | | 1 | 20 | 21 | |

TABLE 2
Tennessee Tutoring Evaluation 2005–2006: Summary of Students Receiving Tutoring Used in the Analysis

| District Name | Provider | Subject | Grade (N) | | | | | Total (N) |
|-----------------|---|---------|-----------|----|----|----|----|-----------|
| | | | 4 | 5 | 6 | 7 | 8 | |
| Davidson County | A to Z In-Home Tutoring | Math | 3 | 9 | 9 | 1 | 4 | 26 |
| | Education Station | Reading | 3 | 2 | 6 | | 4 | 15 |
| | Education Station | Math | 18 | 9 | 14 | 4 | 3 | 48 |
| Hamilton County | A to Z In-Home Tutoring | Math | 17 | 9 | 2 | 2 | 1 | 31 |
| | HONORS Learning Center | Reading | 16 | 8 | 2 | 2 | 1 | 29 |
| | Newton Learning, a Division of Edison Schools | Math | 4 | 4 | 5 | 5 | 4 | 22 |
| | Newton Learning, a Division of Edison Schools | Reading | 4 | 4 | 5 | 5 | 4 | 22 |
| | The Urban League Street Academy | Math | 8 | 18 | 44 | 8 | 12 | 90 |
| | The Urban League Street Academy | Reading | 8 | 18 | 44 | 8 | 11 | 89 |
| | The Urban League Street Academy | Math | 74 | 84 | | | | 158 |
| Memphis | Club Z! In-Home Tutoring | Math | 1 | 1 | 2 | 8 | 3 | 15 |
| | Education Station | Reading | 3 | 4 | 2 | 7 | 5 | 21 |
| | If I Had a Hammer, Inc. | Reading | 6 | 9 | 2 | | | 17 |
| | Innovative Ideas, Inc. | Math | 1 | 2 | 27 | 33 | 47 | 110 |
| | PLATO Learning, Inc. | Reading | 7 | 5 | 9 | | | 21 |
| | X-CEL Operating Foundation | Reading | 10 | 3 | 8 | | | 21 |
| | Innovative Ideas, Inc. | Math | | | 27 | 41 | 21 | 89 |
| | Innovative Ideas, Inc. | Reading | | | 25 | 37 | 26 | 88 |
| | PLATO Learning, Inc. | Math | 7 | 5 | 9 | | | 21 |
| | X-CEL Operating Foundation | Reading | 10 | 3 | 8 | | | 21 |

Note. Students must have predicted scores and teacher linkages to be included in the analyses.

TABLE 3
Tennessee Tutoring Evaluation 2005–2006: Summary of Students Used in the Matched-Pairs Analysis

| District Name | Provider | Subject | N Total | N | | % Matched |
|---|---------------------------------|---------|---------|---------|-----------|-----------|
| | | | | Matched | % Matched | |
| Davidson County | A to Z in-Home Tutoring | Math | 26 | 26 | 100.0 | |
| | | Reading | 15 | 15 | 100.0 | |
| Hamilton County | Education Station | Math | 48 | 46 | 95.8 | |
| | | Reading | 29 | 25 | 86.2 | |
| | A to Z In-Home Tutoring | Math | 31 | 26 | 83.9 | |
| | | Reading | 29 | 25 | 86.2 | |
| | HONORS Learning Center | Math | 22 | 20 | 90.9 | |
| | | Reading | 22 | 18 | 81.8 | |
| Newton Learning, a Division of Edison Schools | Math | Math | 90 | 70 | 77.8 | |
| | | Reading | 89 | 76 | 85.4 | |
| | The Urban League Street Academy | Math | 158 | 105 | 66.5 | |
| | | Reading | 156 | 104 | 66.7 | |
| Memphis | Club Z! In-Home Tutoring | Math | 15 | 15 | 100.0 | |
| | | Reading | 21 | 21 | 100.0 | |
| | Education Station | Reading | 17 | 16 | 94.1 | |
| | | Math | 110 | 102 | 92.7 | |
| | Innovative Ideas, Inc. | Reading | 52 | 50 | 96.2 | |
| | | Math | 89 | 83 | 93.3 | |
| Memphis | PLATO Learning, Inc. | Math | 21 | 20 | 95.2 | |
| | | Reading | 21 | 20 | 95.2 | |
| | X-CEL Operating Foundation | Reading | 20 | 19 | 95.0 | |
| | | Math | 20 | 19 | 95.0 | |

match percentage rates were relatively high, reaching 100% in 4 out of 21 cases (19%), and 90% or higher in 10 cases (48%).

Achievement Outcomes

The results of the statistical analyses for the three models are summarized by provider for Math and R/LA in Tables 4 and 5, respectively. The top portion of the tables indicates, for each model, the significance of (a) the contribution of the predictor variables to the State NCE scores, (b) differences in State NCE scores between the grade levels examined, and (d) differences in NCE scores between participating providers. Although obtaining a highly significant predictor effect is both desirable and expected, there was no particular interest in, or rationale for, finding differences between grade levels or providers. Such differences may occur in given years due to true effects (e.g., tutoring is more effective in certain grades or by certain providers) or to extraneous variables involving, for example, student

TABLE 4
Tennessee Tutoring Evaluation 2005–2006: Analyses of Math in Grades 4–8
Response Variable = Scale Score (in State Units)

| Effect | Model 1 | | | | | | Model 2 | | | | | | Model 3 | | | | | | | | |
|--|--|-------|---------|---|-----------|------|---|---------|---------|---|-----------|-------|---|-------|---------|-----------|-----------|-------|---------|-----------|-------|
| | Fixed Effects: Predicted Score, Grade, Provider | | | Fixed Effects: Predicted Score, Grade, Provider Random: Teacher nested within Grade | | | Fixed Effects: Predicted Score, Grade, Provider Random: Teacher nested within Grade | | | Fixed Effects: Grade, Provider Random Effect: Block (pairs matched by predicted score within teacher) | | | Fixed Effects: Grade, Provider Random Effect: Block (pairs matched by predicted score within teacher) | | | | | | | | |
| | NumDF | DenDF | F-Value | p-Value | Eff. Size | SE | NumDF | DenDF | F-Value | p-Value | Eff. Size | SE | NumDF | DenDF | F-Value | p-Value | Eff. Size | SE | | | |
| Predicted Score | 1 | 15242 | 22376.2 | 0.000 | 0.000 | 1 | 14862 | 21819.4 | 0.000 | 0.000 | 1 | 14862 | 21819.4 | 0.000 | 0.000 | 0.000 | 1 | 14862 | 21819.4 | 0.000 | 0.000 |
| Grade | 4 | 15242 | 3.87 | 0.004 | 0.004 | 4 | 380 | 0.58 | 0.678 | 0.58 | 4 | 552 | 2.48 | 0.043 | 0.043 | 0.043 | 4 | 552 | 2.48 | 0.043 | 0.043 |
| Provider | 10 | 15242 | 1.22 | 0.269 | 0.269 | 10 | 14862 | 0.72 | 0.702 | 0.72 | 10 | 542 | 1.11 | 0.354 | 0.354 | 0.354 | 10 | 542 | 1.11 | 0.354 | 0.354 |
| Comparisons | Est. | SE | p-value | Eff. Size | Est. | SE | Est. | SE | p-value | Eff. Size | Est. | SE | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size | |
| A's and B's Tutorial Service and Learning Institute versus Control | -1.58 | 2.03 | 0.438 | -0.14 | -1.61 | 1.88 | -1.61 | 1.88 | 0.391 | -0.16 | -4.32 | 2.23 | -4.32 | 2.23 | 0.054 | -0.45 | -4.32 | 2.23 | 0.054 | -0.45 | |
| A to Z In-Home Tutoring versus Control | 1.59 | 1.48 | 0.284 | 0.14 | 1.16 | 1.41 | 1.16 | 1.41 | 0.411 | 0.11 | -0.60 | 1.72 | -0.60 | 1.72 | 0.728 | -0.06 | -0.60 | 1.72 | 0.728 | -0.06 | |
| Club Z! In-Home Tutoring versus Control | 3.52 | 2.55 | 0.167 | 0.32 | 2.74 | 2.36 | 2.74 | 2.36 | 0.245 | 0.27 | 0.43 | 2.79 | 0.43 | 2.79 | 0.877 | 0.04 | 0.43 | 2.79 | 0.877 | 0.04 | |
| Education Station versus Control | -0.79 | 1.44 | 0.586 | -0.07 | -1.63 | 1.38 | -1.63 | 1.38 | 0.238 | -0.16 | -3.25 | 1.64 | -3.25 | 1.64 | 0.048 | -0.34 | -3.25 | 1.64 | 0.048 | -0.34 | |
| HONORS Learning versus Control | 0.18 | 2.37 | 0.940 | 0.02 | -0.57 | 2.22 | -0.57 | 2.22 | 0.796 | -0.06 | -4.12 | 2.74 | -4.12 | 2.74 | 0.134 | -0.43 | -4.12 | 2.74 | 0.134 | -0.43 | |

(continued)

TABLE 4 (Continued)

| Comparisons | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size |
|--|----------|----------|----------|-----------|----------|----------|----------|-----------|----------|----------|----------|-----------|
| If I Had a Hammer versus Control | 1.43 | 1.06 | 0.178 | 0.13 | 0.31 | 1.01 | 0.760 | 0.03 | 0.61 | 1.25 | 0.628 | 0.06 |
| Innovative Ideas, Inc. versus Control | 0.98 | 1.18 | 0.407 | 0.09 | 0.23 | 1.14 | 0.840 | 0.02 | -0.79 | 1.39 | 0.571 | -0.08 |
| Newton Learning versus Control | 1.19 | 1.18 | 0.313 | 0.11 | 0.42 | 1.20 | 0.724 | 0.04 | 0.27 | 1.53 | 0.861 | 0.03 |
| PLATO Learning versus Control | -1.35 | 2.43 | 0.579 | -0.12 | 0.74 | 2.37 | 0.756 | 0.07 | -0.76 | 2.77 | 0.784 | -0.08 |
| The Urban League Street Academy versus Control | 1.94 | 0.91 | 0.033 | 0.17 | 1.62 | 0.97 | 0.096 | 0.16 | 0.76 | 1.25 | 0.544 | 0.08 |
| Variance Component Estimates | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate |
| Grade 4 | | | | | 51.15 | | | | | | | 195.62 |
| Grade 5 | | | | | 30.99 | | | | | | | 107.98 |
| Grade 6 | | | | | 22.28 | | | | | | | 176.95 |
| Grade 7 | | | | | 18.35 | | | | | | | 97.15 |
| Grade 8 | | | | | 8.77 | | | | | | | 194.61 |
| Residual | | | | | 123.53 | | | | | | | 92.74 |

Note. Den = denominator.

TABLE 5
Tennessee Tutoring Evaluation 2005–2006: Analysis for Reading/Language Arts in Grades 4–8

| Effect | Response Variable=Scale Score (In State Units) | | | | | | | | | | | |
|--|--|-------|---------|-----------|---------|-------|---------|-----------|---------|-------|---------|-----------|
| | Model 1 | | | | Model 2 | | | | Model 3 | | | |
| | NumDF | DenDF | F-Value | p-Value | NumDF | DenDF | F-Value | p-Value | NumDF | DenDF | F-Value | p-Value |
| Predicted Score | 1 | 12822 | 20056.6 | 0.000 | 1 | 12436 | 16648.9 | 0.000 | 1 | 12436 | 16648.9 | 0.000 |
| Grade | 4 | 12822 | 4.73 | 0.001 | 4 | 386 | 2.48 | 0.044 | 4 | 453 | 3.50 | 0.008 |
| Provider | 10 | 12822 | 2.70 | 0.003 | 10 | 12436 | 1.75 | 0.064 | 10 | 440 | 2.05 | 0.027 |
| Comparisons | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size |
| A to Z In-Home Tutoring versus Control versus Control | -0.65 | 1.56 | 0.677 | -0.06 | 0.19 | 1.55 | 0.904 | 0.02 | 2.43 | 1.95 | 0.214 | 0.24 |
| Club Z! In-Home Tutoring versus Control | 0.55 | 1.98 | 0.783 | 0.05 | 0.63 | 1.96 | 0.749 | 0.06 | 0.15 | 2.43 | 0.951 | 0.01 |
| Education Station versus Control | -0.11 | 2.15 | 0.959 | -0.01 | -0.57 | 2.15 | 0.792 | -0.06 | 0.59 | 2.62 | 0.822 | 0.06 |
| HONORS Learning Center versus Control | 4.10 | 2.19 | 0.062 | 0.40 | 2.63 | 2.16 | 0.225 | 0.26 | 0.72 | 2.80 | 0.797 | 0.07 |
| If I Had a Hammer, Inc. versus Control | 1.13 | 1.43 | 0.430 | 0.11 | 1.06 | 1.42 | 0.457 | 0.11 | 0.69 | 1.76 | 0.693 | 0.07 |

(continued)

TABLE 5 (Continued)

| Comparisons | Response Variable=Scale Score (In State Units) | | | | | | | | | | | |
|--|--|------|---------|-----------|----------|------|---------|-----------|----------|------|---------|-----------|
| | Model 1 | | | | Model 2 | | | | Model 3 | | | |
| | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size | Est. | SE | p-value | Eff. Size |
| Innovative Ideas, Inc. versus Control | 1.08 | 1.10 | 0.328 | 0.10 | 1.81 | 1.12 | 0.104 | 0.18 | 2.31 | 1.43 | 0.108 | 0.23 |
| Newton Learning versus Control | 2.98 | 1.10 | 0.007 | 0.29 | 1.44 | 1.17 | 0.216 | 0.14 | -0.24 | 1.46 | 0.869 | -0.02 |
| PLATO Learning versus Control | -7.42 | 2.25 | 0.001 | -0.72 | -6.82 | 2.27 | 0.003 | -0.68 | -9.86 | 2.71 | 0.000 | -0.99 |
| The Urban League Street Academy versus Control | -0.17 | 0.85 | 0.837 | -0.02 | -0.12 | 0.89 | 0.896 | -0.01 | -0.59 | 1.29 | 0.649 | -0.06 |
| X-CEL Operating Academy versus Control | 4.19 | 2.30 | 0.069 | 0.41 | 3.54 | 2.38 | 0.136 | 0.35 | 4.75 | 2.90 | 0.103 | 0.48 |
| Variance Component Estimates | Estimate | | | | Estimate | | | | Estimate | | | |
| Grade 4 | 105.66 | | | | 4.65 | | | | 62.93 | | | |
| Grade 5 | | | | | 8.22 | | | | 150.40 | | | |
| Grade 6 | | | | | 11.99 | | | | 97.04 | | | |
| Grade 7 | | | | | 2.73 | | | | 82.34 | | | |
| Grade 8 | | | | | 4.09 | | | | 122.04 | | | |
| Residual | | | | | 99.76 | | | | 98.59 | | | |

Note. Den = denominator.

sampling, district characteristics, or testing factors. Also, there was no interest in this study in directly comparing individual providers. The bottom portion of the tables, labeled "Variance Component Estimates," presents estimates of the teacher variance within grades for the two models that controlled for teacher effects. In general, but not consistently, such variance is higher in lower grades.

Most critical to the present evaluation, the "Comparisons" section of the tables presents the results of the statistical comparison of each provider with the control group (the nontutored students) for each model. The column labeled "Est." contains the estimate of the provider effect—the average difference in scores between the students tutored by the provider and the nontutored students. A positive effect means that the students tutored by the provider scored higher, on average, than nontutored students. The column labeled "SE" contains the standard error of the estimate. The column labeled "*p*-value" contains the statistical level of significance for the comparison (an alpha level of .05 was used to determine significance). The column labeled "Eff. Size" contains the estimate of the effect size, defined to be the effect estimate divided by the square root of the residual error.¹ Note that although several of the effect sizes in Math and R/LA (Tables 4 and 5) are moderately large (exceeding +.25 in absolute value), few are statistically significant due to small sample size. Also, the results for the three models are mostly consistent. Where differences are indicated, more credibility can be ascribed to results for Models 2 and 3, which controlled for teacher effects on student achievement, than to those for Model 1.

As may be seen in the top portion of Table 4, the predicted score accounted for significant variance in Math achievement, as expected, for both Models 1 and 2 (Model 3 controlled for predicted scores via the matched student pairings). Grade-level differences were significant in Models 1 ($p = .004$) and 3 ($p = .043$), but not Model 2. Provider differences were not significant in any of the three models.

The "Comparisons" (middle) section of Table 4 shows one statistically significant ($p < .05$) effect in Math for Model 1: The Urban League Street Academy's services were associated with a positive gain in Math of 1.94 NCE points ($p = .033$, Eff. Size = +0.17). For the more rigorous designs controlling teacher effects, no significant differences were indicated in the Model 2 analyses, and one effect was significant in the Model 3 analyses. The latter showed an estimated negative impact of -3.25 NCE points associated with Education Station ($p = .048$, Eff. Size = -0.34). Results for individual providers, however, need to be interpreted very cautiously given the relatively low sample sizes and potential confounding with

¹The effect size represents the number of standard deviations by which the SES student achievement mean differed from the nontutored student mean. Generally, effect sizes of +0.20 or higher (indicating an advantage for the treatment group of at least one-fifth SD) are considered to be moderate to strong in impact, and educationally meaningful.

school and environmental factors. Replication of effects across several years would provide increased credence.

Results for R/LA (Table 5) revealed significance for predicted scores in Models 1 and 2, grade-level differences in all three models, and Provider differences in Models 1 and 3. More critically, the analyses of provider effects revealed several significant outcomes. For the Model 1 analyses, there was a significant positive effect for Newton Learning and a negative effect for PLATO Learning. Controlling for teacher effects, Model 2 analyses, however, yielded only one significant effect—a negative impact of -6.82 points for PLATO Learning ($p = .003$, Eff. Size = -0.68). The latter effect for the same provider was corroborated by the Model 3 (matched pair) analysis, which yielded an estimated deficit of -9.86 NCE points ($p < .001$, Eff. Size = -0.99).

Participant/Stakeholder Perceptions

Results from the questionnaires administered to the four stakeholder/participant groupings are presented in the following sections.

District coordinator. A District Coordinator Questionnaire was completed individually for each provider serving in the three SES districts. Aggregate results are presented in Table 6 and summarized briefly in the following. Open-ended comments were relatively few and not informative for evaluation purposes. In total, the respondents completed 39 separate surveys. As shown in the upper section of Table 6, the coordinators clearly responded positively about the frequency with which providers communicated with them (94.9% frequently or occasionally) and met their obligations (89.7%). By comparison, the providers were viewed as less, but still somewhat (close to 50% frequently or occasionally), active in communicating with teachers and parents, and collaborating with them to set goals for students. Interestingly, these results contrast somewhat with those from the one Jefferson County Public Schools (Kentucky) coordinator surveyed by Muñoz et al. (this issue), who viewed the providers as more active with teachers and parents than with him.

The lower section (Likert-type ratings) of Table 6 further reflects mostly positive reactions to the provider services, especially regarding compliance with federal and state/local laws (over 90% agreement on both). There was less agreement (lower than 50%) that providers adapted the tutoring to the school's curriculum or integrated it with classroom learning activities. Of note, the district coordinators expressed overall satisfaction with providers and with the providers' effectiveness in raising achievement (close to 75% agreement on both).

School leader questionnaire. The school leader questionnaire was completed by 50 respondents, 9 of whom identified themselves as principals or assis-

TABLE 6
Responses to the District Coordinator Survey (n = 39)

| Items | % Responding | | | | |
|--|----------------|--------------|------------|------------|-------------------|
| | Frequently | Occasionally | Not at All | Don't Know | |
| How often does the provider... | | | | | |
| 1. Communicate with you during the year? | 43.6 | 51.3 | 5.1 | 0.0 | |
| 2. Meet the obligations for conducting tutoring sessions? | 56.4 | 33.3 | 2.6 | 7.7 | |
| 3. Communicate with teachers during the school year? | 10.3 | 33.3 | 5.1 | 51.3 | |
| 4. Communicate with parents during the year? | 17.9 | 38.5 | 2.6 | 41.0 | |
| 5. Collaborate with you to set goals for student growth? | 7.7 | 43.6 | 48.7 | 0.0 | |
| | Strongly Agree | Agree | Undecided | Disagree | Strongly Disagree |
| The provider... | | | | | |
| 6. Adapted the tutoring services to this school's curriculum. | 5.1 | 43.6 | 38.5 | 12.8 | 0.0 |
| 7. Integrated the tutoring services with classroom learning activities. | 2.6 | 43.6 | 41.0 | 12.8 | 0.0 |
| 8. Aligned their services with state and local standards. | 2.6 | 87.2 | 7.7 | 2.6 | 0.0 |
| 9. Offered services to Special Education and ELL students. | 0.0 | 56.4 | 33.3 | 10.3 | 0.0 |
| 10. Complied with applicable federal NCLB laws. | 28.2 | 64.1 | 2.6 | 5.1 | 0.0 |
| 11. Complied with applicable state and local laws. | 28.2 | 61.5 | 2.6 | 5.1 | 0.0 |
| Overall Assessment: | | | | | |
| 12. I believe the services offered by this provider positively impacted student achievement. | 7.7 | 66.7 | 20.5 | 5.1 | 0.0 |
| 13. Overall, I am satisfied with the services of this provider. | 5.1 | 69.2 | 20.5 | 5.1 | 0.0 |

Note. Some percentages may not add up to 100% due to not all parents answering the question.

tant principals. The remainder were site coordinators, some full-time and others combining the duties with regular teaching responsibilities. Approximately one-fourth (28%) of the respondents indicated that they were employed by the provider in some capacity. Table 7 summarizes the closed-ended responses.

The majority of the respondents were positive about provider communications, perceiving such to be frequent 36% of the time with them (the school leaders), 60%

TABLE 7
Responses to Principal/School Leader Survey (n = 50)

| Items | % Responding | | | | |
|--|----------------|--------------|------------|------------|-------------------|
| | Frequently | Occasionally | Not at All | Don't Know | |
| How often does the provider... | | | | | |
| 1. Communicate with you during the year? | 36.0 | 32.0 | 20.0 | 12.0 | |
| 2. Meet the obligations for conducting tutoring sessions? | 58.0 | 16.0 | 6.0 | 20.0 | |
| 3. Communicate with teachers during the school year? | 60.0 | 24.0 | 16.0 | 0.0 | |
| 4. Communicate with parents during the year? | 42.0 | 24.0 | 8.0 | 26.0 | |
| | Strongly Agree | Agree | Undecided | Disagree | Strongly Disagree |
| The provider... | | | | | |
| 5. Adapted the tutoring services to this school's curriculum. | 34.0 | 24.0 | 28.0 | 10.0 | 4.0 |
| 6. Integrated the tutoring services with classroom learning activities. | 30.0 | 24.0 | 26.0 | 12.0 | 6.0 |
| 7. Aligned their services with state and local standards. | 40.0 | 22.0 | 34.0 | 0.0 | 4.0 |
| 8. Offered services to Special Education and ELL students. | 36.0 | 20.0 | 36.0 | 4.0 | 4.0 |
| 9. Complied with applicable federal, state and local laws. | 36.0 | 30.0 | 26.0 | 4.0 | 2.0 |
| Overall Assessment: | | | | | |
| 10. I believe the services offered by this provider positively impacted student achievement. | 36.0 | 22.0 | 30.0 | 6.0 | 6.0 |
| 11. Overall, I am satisfied with the services of this provider. | 36.0 | 22.0 | 30.0 | 6.0 | 6.0 |

Note. Some percentages may not add up to 100% due to not all parents answering the question.

of the time with teachers, and 40% with parents. Providers were also viewed by the majority (58%) as frequently meeting their tutoring obligations. Compared to the district coordinators, the school leaders more frequently agreed (from 54%–66%) that the providers (a) adapted tutoring to the classroom, state, and local standards; (b) accommodated special education and English language learner students; and (c) complied with policies and laws. Two-thirds believed both that the tutoring services positively impacted student achievement and that they were satisfied overall with the provider.

Of the 19 open-ended comments that were submitted, approximately one-third each were positive, negative, and neutral. Positive comments consisted of brief superlative remarks (e.g., “excellent program,” “good provider”) or, in a few cases, longer and more specific impressions such as, “This provider was very efficient and accurate. Our students gained knowledge and expertise on enhancing their math skills.” Negative comments focused on lack of communication (e.g., “Administrators only received one comment from the provider concerning these students”) and disorganization (e.g., “The services were unorganized from the initial set up to the implementation of the program”). Neutral comments predominantly expressed lack of contact with or information about the providers’ services.

Teacher questionnaire. Responses by the teachers ($n = 128$) in the SES schools are summarized in Table 8. Of those responding, 28% indicated that they were employed by the provider. A comparison of responses by that group and the remaining teachers indicated no discernable difference (i.e., those with and without ties to a provider appeared equally likely to voice criticism or praise).

Teachers’ reactions to the provider services, overall, were positive and consistent with those just reviewed for the school leaders. Notably, about one-third (33.6%) of the respondents reported frequent communications by the provider, and another one-third (32%) reported occasional communications. A slight majority (51.6%) believed that the providers frequently met their obligations for conducting tutoring sessions. As might be expected if firsthand knowledge were lacking, teachers were less certain that providers frequently communicated with parents and principals (see Table 8). Also encouraging were the teachers’ relatively high levels of agreement (about 70% or higher) that providers adapted services to the curriculum, integrated the tutoring with classroom learning, aligned the services with the state and local standards, offered services to special education students, and complied with laws and policies. Almost three-fourths of the respondents, a higher percentage than the school leaders (Table 7), agreed both that the tutoring services positively impacted student achievement, and that they were satisfied overall with the provider.

Open-ended comments were mostly positive (43%) or neutral (38%), with the latter expressing little or no knowledge that their students were receiving tutoring. The few negative comments (19%) conveyed two main themes—lack of timeliness

TABLE 8
Teacher Questionnaire Responses ($n = 128$)

| Items | % Responding | | | | |
|--|----------------|--------------|------------|------------|-------------------|
| | Frequently | Occasionally | Not at All | Don't Know | |
| How often does the provider... | | | | | |
| 1. Communicate with you during the school year? | 33.6 | 32.0 | 21.1 | 13.3 | |
| 2. Meet the obligations for conducting tutoring sessions? | 51.6 | 20.3 | 0.8 | 26.6 | |
| 3. Communicate with the principal or school site coordinator during the year? | 43.8 | 18.8 | 2.3 | 35.2 | |
| 4. Communicate with parents during the year? | 29.7 | 33.6 | 2.3 | 34.4 | |
| | Strongly Agree | Agree | Undecided | Disagree | Strongly Disagree |
| The provider... | | | | | |
| 5. Adapted the tutoring services to this school's curriculum. | 29.7 | 39.8 | 25.8 | 4.7 | 0.0 |
| 6. Integrated the tutoring services with classroom learning activities. | 28.1 | 39.8 | 25.0 | 7.0 | 0.0 |
| 7. Aligned their services with federal, state and local standards. | 36.7 | 34.4 | 26.6 | 2.3 | 0.0 |
| 8. Offered services to Special Education and ELL students. | 24.2 | 34.4 | 37.5 | 3.1 | 0.8 |
| 9. Complied with applicable federal, state and local laws. | 33.6 | 35.9 | 30.5 | 0.0 | 0.0 |
| Overall Assessment: | | | | | |
| 10. I believe the services offered by this provider positively impacted student achievement. | 35.9 | 38.3 | 19.5 | 3.9 | 2.3 |
| 11. Overall, I am satisfied with the services of this provider. | 34.4 | 33.6 | 23.4 | 3.9 | 3.9 |

Note. Some percentages may not add up to 100% due to not all parents answering the question.

in providing tutoring and poor adaptation to students' grade levels or abilities. For example, a dissatisfied teacher conveyed, "Grades were combined; the program was geared toward the higher grade and most of the children enrolled were already having difficulty. I don't think it met their needs." The positive comments tended to be general in theme, indicating satisfaction with the provider and the hope that the tutoring would continue the next year. For example, one teacher wrote, "I truly think the program has helped our students." Another indicated, "On-site provider is well trained and conscientious."

Parent questionnaire. Parent reactions to the tutoring ($n = 355$), predictably, were the most positive of all respondent groups. As shown by the response summary in Table 9, from about 80% to 90% of the parents indicated that the providers communicated with them frequently, or at least occasionally. Two-thirds believed that the tutors frequently helped their child with material being taught in the regular classroom; almost 80% indicated that the tutoring sessions frequently started and ended on time. Impressively, over 90% of the parents agreed that the tutoring services positively impacted student achievement, and that they were satisfied overall with the provider. The respondents voiced comparable high levels of agreement regarding the district's activities in providing information about the availability of SES and helping them to select appropriate providers.

Open-ended comments by parents ($n = 34$) were predominately positive (76.5%). The most salient theme emphasized the help that children were receiving for academic improvement. Exemplary comments were:

- "This program has helped my child with his reading. Thank you."
- "My child was helped a great deal with the one-on-one tutoring this center gave. I do hope it is available to him in the coming years."
- Negative comments (20.5%) stressed two themes—lack of communication on student progress and insufficient tutoring time. Exemplary comments are:
- "The provider tutored my child for 14 hrs instead of 27 hrs. Somebody has to follow up with the services."
- "Please let me know of my child's progress. I haven't received any progress reports or anything."

Cross-group comparison. The responses of the four groups to comparable questionnaire items on their respective surveys are summarized in Table 10. As shown, provider communications were perceived to be more frequent by district coordinators and parents than by teachers and site coordinators. Teachers were the most favorable regarding the connection between the tutoring and the school's curriculum and classroom learning activities, whereas district coordinators were most supportive of the providers' compliance with laws and alignment to standards. Parents were clearly most positive and site coordinators least posi-

TABLE 9
Parent Questionnaire Responses in Tennessee 2005-2006 ($n = 355$)

| Parent Questionnaire Item | Percent | | | | |
|---|----------------|--------------|------------|------------|-------------------|
| | Frequently | Occasionally | Not at All | Don't Know | |
| How often does the provider... | | | | | |
| 1. Talk to me about my child's progress? | 52.1 | 30.1 | 13.2 | 3.1 | |
| 2. Talk to my child's teachers about his/her progress? | 48.7 | 28.5 | 6.5 | 14.9 | |
| 3. Send letters or notes home to me about my child's progress? | 43.7 | 36.6 | 14.6 | 2.8 | |
| 4. Help my child with subjects they are working on in their classroom at school? | 67.6 | 23.1 | 2.5 | 5.4 | |
| 5. Start and end the tutoring sessions at the scheduled time? | 78.0 | 11.8 | 2.8 | 5.1 | |
| Parent Questionnaire Item | Percent | | | | |
| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree |
| Overall assessment: | | | | | |
| 6. I believe that the services offered have helped my child's achievement. | 63.9 | 26.2 | 7.6 | 0.6 | 1.1 |
| 7. Overall, I am pleased with the services that my child received. | 64.2 | 28.2 | 4.8 | 0.0 | 1.4 |
| District assessment: | | | | | |
| 8. I was given information about my child's rights under the No Child Left Behind law. | 60.3 | 30.7 | 3.4 | 2.5 | 1.7 |
| 9. I was given enough time to decide which service provider I wanted for my child. | 53.0 | 34.9 | 6.2 | 2.5 | 2.0 |
| 10. I am pleased with the way my school district helped me obtain Supplemental Educational Services for my child. | 62.5 | 29.0 | 4.8 | 1.7 | 1.4 |

Note. Some percentages may not add up to 100% due to not all parents answering the question.

TABLE 10
Cross-Group Comparison of Questionnaire Responses

| Questionnaire Item | Percent Frequently and Occasionally | | | |
|---|-------------------------------------|-----------------------|---|----------------------|
| | Parents (n = 355) | Teachers (n = 128) | Principal/Site Coordinator (n = 50) | District (n = 39) |
| How often does the provider... | | | | |
| 1. Communicate with me during the school year? | 82.2 | 65.6 | 68.0 | 94.9 |
| 2. Meet the obligations for conducting tutoring services (start/end on time)? | 89.8 | 71.9 | 74.0 | 89.7 |
| The provider... | | | | |
| 3. Adapted the tutoring services to this school's curriculum. | n/a | 69.5 | 58.0 | 48.7 |
| 4. Integrated the tutoring services with classroom learning activities | n/a | 67.9 | 54.0 | 46.2 |
| 5. Aligned their services with state and local standards. | n/a | 71.1 | 62.0 | 89.8 |
| 6. Offered services to Special Education and ESL students. | n/a | 58.6 | 56.0 | 56.4 |
| 7. Complied with applicable federal, state, and local laws. | n/a | 69.5 | 66.0 | 92.3 |
| Overall assessment: | | | | |
| 8. I believe that the services offered have helped my child's achievement. | 90.1 | 74.2 | 58.0 | 74.4 |
| 9. Overall, I am pleased with the services that my child received. | 92.4 | 68.0 | 58.0 | 74.3 |
| District assessment: | | | | |
| 10. I was given information about my child's rights under the No Child Left Behind law. | 91.0 | n/a | n/a | n/a |
| 11. I was given enough time to decide which service provider I wanted for my child. | 87.9 | n/a | n/a | n/a |
| 12. I am pleased with the way my school district helped me obtain Supplemental Educational Services for my child. | 91.5 | n/a | n/a | n/a |

tive regarding the effectiveness of the tutoring in general and for improving achievement in particular.

DISCUSSION

These results, overall, show generally positive reactions toward the SES tutoring and specific providers by district coordinators, school leaders, teachers, and, especially, parents. However, the effects of tutoring on student achievement tended to be small and, with only a few exceptions, not significantly different from zero. These findings are consistent with those of most prior studies (Muñoz et al., this issue; Office of Research, Evaluation, and Accountability, 2007; Potter et al., 2006; Rickles & Barnhart, 2007; Rickles & White, 2006), but are slightly more positive with regard to stakeholder perceptions and less supportive of SES effects on student achievement. For example, 71.9% of the teachers in the Jefferson County Public Schools (JCPS) study (Muñoz et al., this issue), as compared to only 21.1% of those in this study, responded that the provider communicated with them "not at all." The JCPS district coordinator indicated that 93% of the providers failed to work with him at all to set goals for student growth, whereas only 50% of Tennessee coordinators perceived that to be the case. The other half of the Tennessee coordinators rated such collaboration as occurring frequently or occasionally. Regarding student achievement, although the Kentucky study (Muñoz et al., this issue) also failed to find SES impacts on student achievement, other studies (Rickles & White, 2006; Zimmer et al., 2006; Zimmer et al., 2007) have detected some significant effects, albeit small in magnitude.

One underlying assumption of this study is that, even though theoretically and logically 30–40 hrs of after-school tutoring are not likely to produce strong effect sizes (particularly on high-stakes standardized assessments), it is both practically and scientifically important to conduct such analyses in the absence of existing evidence to the contrary. If results from multiple rigorous studies, in fact, did indicate consistent, meaningful benefits (say, effect sizes exceeding 0.20), the implications for policy and practice would be to promote strongly continuance of SES as a core NCLB (2001; or school reform) intervention. However, in the absence of such effects, the implications are less clear. There is much logical and empirical justification for believing that one-on-one tutoring from qualified adults does improve student learning and raises achievement (Slavin, 2006; Tingley, 2001). But what remains to be determined is whether or not SES, as a unique systemic and nationally-based program, is typically implemented in a manner likely to produce such benefits, and if so, the associated impacts on standardized AYP test performance in particular.

With regard to the implementation of SES, only indirect evidence reflecting the views of various participant and stakeholder groups was available in this study.

The generally positive responses indicate support by respondents for the activities of the SES providers and quality of the tutoring services. However, given the voluntary nature of participation in the surveys, the representativeness of the respondents in each stakeholder group certainly can be questioned. Based on our contacts with many members of these groups during the 4-year duration of the Tennessee state SES evaluations, we strongly suspect a sampling bias in which survey responders were more likely than nonresponders to be involved in and positive about SES. In particular, our impressions from these experiences are that provider communications with teachers and their integration of tutoring into regular classroom teaching were much less frequent than what the respondent sample indicated and, therefore, closer to the findings of the Kentucky study (Muñoz et al., this issue). Generic types of tutoring not directly linked to everyday classroom learning may be poorly adapted to student needs (Slavin & Fashola, 1998). As Muñoz et al. (this issue) point out, it's even conceivable that such tutoring could cause interference by teaching content or skills in ways that differ from those emphasized by the regular classroom teacher.

Even if the SES tutoring were highly adaptive to student needs, a point already raised several times in this article is that 30–40 total hr of one-on-one tutoring may have, at best, only limited potential to raise student achievement on standardized tests. Obviously, there are so many factors, such as teacher effectiveness (Ross et al., 2003), school leadership (Datnow, Hubbard, & Mehan, 2002), and interventions such as comprehensive school reform, extended-day, reduced class size (Desimone, 2002; Slavin & Fashola, 1998), or special reading and math programs (e.g., Correnti & Rowan, 2007) that potentially exert much stronger effects on student learning or (minimally) serve as confounding variables that mask tutoring effects. In the recent national study of SES, encompassing seven school districts (Zimmer et al., 2007), overall effect sizes calculated for mathematics and reading were .09 and .08, respectively. But because use of a randomized experimental design was precluded, how much of this effect is attributable to differential characteristics (e.g., aptitude, motivation, parental involvement) of the SES participants and nonparticipants (the control group) is unclear. Such potential biases can be particularly troublesome for choice programs such as SES, where fewer than 20% of those eligible for the services actually enroll and participate (Sunderman, 2006).

Based on the prior research previously cited, it might be concluded that typical district implementations of SES yield a likely effect size for achievement gains on state assessments somewhere between $d = 0.0$ and $+0.10$. A possible argument is that this range underestimates the true SES effect because the quasi-experiments conducted thus far have lower internal validity and power for detecting tutoring impacts than do randomized designs. A second, related argument is that power is further weakened substantially by the inability to account for the much larger variance in student achievement attributable to teachers, classrooms, and schools. Using the SAS EVAAS methodology in this study, we were able to exert such con-

trols in two models (teachers nested within grades and matched SES-control student pairs), and found more positive than negative but still generally weak, SES effects rarely surpassing $d = +.010$.

Policymakers and educational leaders tend to view achievement scores as the bottom line in judging whether educational interventions are working (Graczewski, Ruffin, Shambaugh, & Therriault, 2007). So, despite the positive perceptions of the surveyed parent, teacher, school, and state participants/stakeholders in Tennessee, the achievement benefits for SES will need to improve over those of this study and prior state studies to maintain public support for the tutoring as a justifiable use of Title I funding. Possible strategies for providers to increase effectiveness include the following. First, connection of the tutoring to regular classroom instruction in accord with students' individual needs should be increased. An essential part of this process will be communicating more regularly with teachers about students' progress, strengths, and weaknesses both in the classroom and in tutoring sessions. Second, to the extent that the requirement for schools to offer SES is linked to their performance on state assessments, providers need to ensure that the tutoring directly addresses state standards and test content. This suggestion is not to imply that tutoring must be oriented to teaching to the test, but rather that tutors should have explicit awareness of the content most likely to be assessed and the testing formats used (i.e., typically multiple choice). Third, in accord with the extensive literature on program evaluation and school reform (Desimone, 2002; Ross, 2007), providers should concentrate on ensuring strong implementations of the tutoring services to maximize positive effects.

This study, and the others in this issue of *JESPAR*, also suggest directions to improve the quality and validity of future state and district evaluations of SES. One recommendation is to supplement state assessments, perhaps in a smaller study involving a subsample of selected providers that are varied each year, with more sensitive measures, such as individually administered tests of reading comprehension or a constructed-response problem-solving test in mathematics. In addition to providing more authentic measures of learning than do multiple-choice items, these tests conceivably would be less influenced by regular school curricula and programs directed toward the high stakes attainment of AYP. A second recommendation, which again might be implemented with subsamples of students and providers, is to conduct smaller, mixed-methods quasi-experiments in which SES outcomes can be interpreted more fully relative to contextual variables (e.g., the curriculum and other interventions) that impact both SES and comparison students. Third, such quasi-experiments should include some observation of actual tutoring sessions (visited randomly and unannounced if feasible) to obtain first-hand impressions of the pedagogy employed, tutor preparation and effectiveness, and student engagement. Fourth, although this study (see also Muñoz et al., this issue) solicited perceptions of different aspects of SES tutoring from multiple stakeholder groups, important voices that were not heard were those from the students

who were tutored. SES may well have important positive impacts on students that extend beyond achievement gains on state assessments. Such benefits might include increased motivation, self-efficacy, metacognitive skills, and development of effective learning habits. These considerations suggest only the potential of SES to support at-risk students. We hope that the reality will be demonstrated from improved SES program design and delivery, coupled with more rigorous and comprehensive research on its effects.

ACKNOWLEDGMENTS

This study was supported by funding provided by the Tennessee Department of Education. We thank Julie McCarger and Carol Groppe from Tennessee Department of Education for their support and assistance in the conduct of the study.

REFERENCES

- Ascher, C. (2006). NCLB Supplemental Educational Services: Is this what our students need? *Phi Delta Kappan*, 88(2), 136–141.
- Borman, G. D., & D'Agostino, J. V. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis*, 18, 309–326.
- Borman, G. D., Hewes, G., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73, 125–230.
- Burch, P., Steinberg, M., & Donovan, J. (2007). Supplemental Educational Services and NCLB: Policy assumptions, market practices, emerging issues. *Educational Evaluation and Policy Analysis*, 29, 115–133.
- Center on Educational Policy. (2006). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Author.
- Chatterji, M. (2005). Evidence on “what works”: An argument for extended-term mixed-method (ETTM) evaluation designs. *Educational Researcher*, 34(5), 14–24.
- Correnti, R., & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal*, 44, 298–339.
- Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. New York: Routledge-Farmer.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72, 433–480.
- Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, 21, 97–109.
- Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review*, 31, 569–583.
- Good, T. L., & Brophy, J. E. (1987). *Looking into classrooms* (4th ed.). New York: Harper & Row.
- Graczewski, C., Ruffin, M., Shambaugh, L., & Therriault, S. B. (2007). Selecting and implementing whole school improvement models: A district and school administrator perspective. *Journal of Education for Students Placed At Risk*, 12, 71–90.
- Green, S. K., Alderman, G., & Liechty, A. (2004). Peertutoring, individualized intervention, and progress monitoring with at-risk second-grade readers. *Preventing School Failure*, 49, 11–17.
- Ismail, H. N., & Alexander, J. M. (2005). Learning within scripted and nonscripted peer-tutoring sessions: The Malaysian context. *The Journal of Educational Research*, 99(2), 66–77.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Kidron, Y., & Darwin, M. J. (2007). A systematic review of whole school improvement models. *Journal of Education for Students Placed At Risk*, 12, 9–36.
- Lauer, P. A., Motoko, A., Wilkerson, S., Aphorp, H., Snow, D., & Martin-Glenn, M. (2003). *The effectiveness of out-of-school-time strategies in assisting low-achieving students in reading and mathematics: A research synthesis*. Aurora, CO: MCREL.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching*. Upper Saddle River, NJ: Pearson Education.
- No Child Left Behind Act. (2001). Pub. L. No. 107–110.
- Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, 37, 123–151.
- Office of Research, Evaluation, and Accountability; Office of Extended Learning Opportunities. (2007). *SES Tutoring Program Evaluation Report—Year 3*. Chicago: Chicago Public Schools.
- Potter, A., Ross, S. M., & McDonald, A. J. (2005a). *The State of Tennessee, Supplemental Educational Services District Coordinator Questionnaire*. Center for Research in Educational Policy, University of Memphis, Memphis, TN.
- Potter, A., Ross, S. M., & McDonald, A. J. (2005b). *The State of Tennessee, Supplemental Educational Services Parent Questionnaire*. Center for Research in Educational Policy, University of Memphis, Memphis, TN.
- Potter, A., Ross, S. M., & McDonald, A. J. (2005c). *The State of Tennessee, Supplemental Educational Services Principal/Site Coordinator Questionnaire*. Center for Research in Educational Policy, University of Memphis, Memphis, TN.
- Potter, A., Ross, S. M., & McDonald, A. J. (2005d). *The State of Tennessee, Supplemental Educational Services Teacher Questionnaire*. Center for Research in Educational Policy, University of Memphis, Memphis, TN.
- Potter, A., Ross, S. M., & McDonald, A. J. (2005e). *Supplemental Educational Services in the State of Tennessee: 2003–2004*. Center for Research in Educational Policy, University of Memphis, Memphis, TN.
- Potter, A., Ross, S. M., McDonald, A. J., Nunnery, J. A., & Paek, J. (2006). *Supplemental Educational Services in the State of Tennessee: 2004–2005*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- Potter, A., Ross, S. M., Paek, J., McKay, D., Pribesh, S., & Nunnery, J. A. (2007). *Supplemental Educational Services for the Commonwealth of Virginia for 2005–2006*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- Potter, A., Ross, S. M., Paek, J., Pribesh, S. L., & Nunnery, J. A. (2006). *Supplemental Educational Services in the State of Louisiana: 2004–2005*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.
- Rickles, J. H., & Barnhart, M. K. (2007). *The impact of supplemental educational services on student achievement: 2005–2006*. Los Angeles, CA: Los Angeles Unified School District Planning, Assessment, and Research Division.
- Rickles, J. H., & White, J. A. (2006). *The impact of supplemental educational services participation on student achievement*. Los Angeles, CA: Los Angeles Unified School District Planning, Assessment, and Research Division.

- Ross, S. M. (2007). Achievements, challenges, and potential advancements in reviewing educational evidence for consumers. *Journal of Education for Students Placed At Risk*, 12, 1–10.
- 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.
- Ross, S. M., Stringfield, S., Sanders, W. L., & Wright, S. P. (2003). Inside systemic elementary school reform: Teacher effects and teacher mobility. *School Effectiveness and School Improvement*, 14, 73–110.
- Sanders, W. L., & Horn, S. P. (1995). The Tennessee Value-Added Assessment System: A quantitative outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid education measure?* (pp. 130–162). Thousand Oaks, CA: Corwin Press.
- Schaps, E. (2007). Why the No Child Left Behind Act is unsalvageable. *Education Week*, 26(36), 32–33.
- Slavin, R. E. (2006). *Educational psychology: Theory and practice* (8th ed.). Boston: Pearson Education.
- Slavin, R. E., & Fashola, O. S. (1998). *Show me the evidence! Proven and promising programs for American schools*. Thousand Oaks, CA: Corwin.
- Sunderman, G. L. (2006). Do Supplemental Educational Services increase opportunities for minority students? *Phi Delta Kappan*, 88, 117–122.
- Teddlie, C., & Reynolds, D. (Eds.). (2000). *The international handbook of school effectiveness research*. London: Falmer.
- Tingley, J. (2001). Volunteer programs: When good intentions are not enough. *Educational Leadership*, 68(7), 53–55.
- U.S. Department of Education. (2006). *National assessment of Title I: Interim report: Executive summary*. Washington, DC: Author.
- Wasik, B. A. (1997). Volunteer tutoring programs: Do we know what works? *Phi Delta Kappan*, 70, 283–287.
- Zimmer, R., Christina, R., Hamilton, L., & Prine, D. W. (2006). *Evaluation of two out-of-school programs in Pittsburgh Public Schools: No Child Left Behind's Supplemental Educational Services and State of Pennsylvania's Educational Assistance Program*. Santa Monica, CA: RAND.
- Zimmer, R., Gill, B., Razquin, P., Booker, K., & Lockwood, J. R., III. (2007). *State and local implementation of the No Child Left Behind Act: Volume I—Title I school choice, Supplemental Educational Services, and student achievement*. Washington, DC, U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Programs Study Service.