

Evaluating Supplemental Educational Service Providers: Suggested Strategies for States

March 2005

An Issue Brief Prepared by

**Dr. Steven M. Ross
Center for Research in Educational Policy
<http://crep.memphis.edu>**

and

**Supplemental Educational Services Quality Center
www.tutorsforkids.org**



I. INTRODUCTION

This Issue Brief is designed to assist state educational agencies (SEAs) in creating an effective system to evaluate state-approved supplemental educational services (SES) providers. It will help readers to determine evaluation measures, identify possible evaluation methodologies, and address the technical and practical considerations associated with an evaluation. Although this brief is of primary interest to states, it can also help school districts and SES providers understand their roles in the evaluation process.

In the process of selecting an evaluation system, states may consider evaluating providers on three performance dimensions: effectiveness, customer satisfaction, and service delivery. This Issue Brief explains these three performance dimensions and provides an overview of possible evaluation measures and designs that states can use to assess these performance dimensions.

Overview of Supplemental Educational Services

SES is a provision of Title I of the Elementary and Secondary Education Act (ESEA), as reauthorized by the *No Child Left Behind Act* (NCLB). 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 three years (i.e., a school that is at least in its second year of “school improvement” status). A variety of organizations are eligible to provide SES, after receiving formal approval from the state:

- For-profit companies.
- Nonprofit groups.
- Local community programs.
- Colleges or universities.

- National organizations.
- Faith-based groups.
- Private and charter schools.
- Public schools and districts (not identified as in need of improvement).

Roles of States and Districts

NCLB and the U.S. Department of Education require states and districts to fulfill distinct roles in SES implementation. Each state is charged with identifying schools that do not make AYP and schools that have been identified as in need of improvement, corrective action, or restructuring. Each state develops its own application and criteria to become a state-approved SES provider.

States must also monitor each approved provider. States are required to evaluate provider effectiveness after at least 2 years, but they may choose to monitor providers more frequently.¹ The standards used to evaluate providers should be consistent for each provider and should be aligned with the criteria set forth in the state's application to be an SES provider. According to NCLB, at a minimum, states must remove providers from the approved list if they fail to:

- Increase students' achievement for 2 consecutive years.
- Provide services consistent with applicable Federal, state, and local health, safety, and civil rights requirements.

States may collect and report additional information about providers, including parent and student satisfaction with services. The SEA may also request that districts help to monitor SES providers.

Districts are charged with determining which students in Title I schools identified as in need of improvement are eligible for SES.² SES is available to low-income students in those schools—generally, those students eligible for free or reduced-price lunch. Once the district determines which students are eligible, it notifies families at least once per year that their children qualify for SES. The district must also provide families with information about local SES providers to help them choose a provider. If families ask for assistance, the district must help them choose an SES provider. If families' demand for SES exceeds the available funds, districts must then give priority to the lowest-achieving eligible students. Once families choose a provider, the district contracts with and pays the provider for its services. When establishing contracts, districts and providers should work together to clearly indicate how student performance will be measured.

¹ States are encouraged to monitor providers at least annually.

² If a Title I school does not make AYP for 2 consecutive years, it must offer all students in that school the opportunity to transfer to another public school that is not in need of improvement. Additional information about the "choice options" provided by NCLB is available from the U.S. Department of Education's Web site: www.ed.gov/about/offices/list/oii/about/choice.html.

WORKING TOGETHER: CONNECTING STATE AND DISTRICT EFFORTS TO IMPROVE EVALUATION

States and districts bear distinct and critical responsibilities for implementing SES. Experience suggests that regular communication and collaboration can help states and districts carry out these responsibilities more effectively. States are responsible for evaluating SES providers. States may ask school districts for assistance, particularly if districts are already gathering data about local providers (e.g., details about each provider's service delivery). In such instances, states should ensure that the district's role is appropriate and unbiased, because school districts that are *not* in need of improvement may also be SES providers. States and districts can also collaborate to ensure that district contracts with providers establish procedures needed by states to gather data and evaluate providers. Additionally, many of the evaluation options discussed in this Brief require that data gathering and reporting mechanisms be consistent statewide. If data collection infrastructures vary across districts, a state may consider implementing a statewide data collection and compilation mechanism. (For more guidance on data collection options, consult "Technology and Database Considerations" box on page 10.)

II. GUIDING QUESTIONS FOR EVALUATING SES PROVIDERS

To effectively monitor SES providers, states must develop a system that will offer valid and highly usable data to evaluate the impact of each SES provider's services. States may consider a provider's impact in three areas:

1. **Effectiveness.** Did the provider increase student achievement in reading/language arts or math?
2. **Customer satisfaction.** Are parents and students who received SES satisfied?
3. **Service delivery.** Did the provider comply with applicable state and district laws and contractual procedures associated with the delivery of SES?

It is important to remember that a provider will need to demonstrate effectiveness across all of the types of students that it serves, including English language learners (ELLs) and students with disabilities and a variety of learning needs.

States may also monitor SES providers' instructional programs to determine whether the services delivered match those described in the providers' applications to the state. Depending on specific interests and resources, states can examine whether providers are meeting expectations with regard to:

- Experience and qualifications of the tutors.
- Amount of tutoring time received by students.
- Teaching strategies used.
- Instructional grouping and student-instructor ratios.
- Communication with teachers and parents.

III. DETERMINING EVALUATION MEASURES

This section outlines possible assessment measures for a comprehensive evaluation and monitoring effort. Both technical and practical considerations are offered for each measure to help states select an evaluation design that meets their needs. Suggestions on how to incorporate these measures into an evaluation design are provided in section “IV. Evaluation Designs: Student Achievement.”

1. Effectiveness Measures

Measures of impact on student academic achievement are critical to a state’s evaluation of SES providers. This is especially true because NCLB requires that, at a minimum, states remove providers from their approved list if the provider fails to increase students’ achievement for 2 consecutive years. State evaluations could measure achievement levels through state-mandated assessments, supplementary individualized assessments, or provider-developed assessments.

a. Student-level test scores from state-mandated assessments

Evaluations of SES providers can examine students’ scores on yearly achievement assessments that are administered by the state in compliance with NCLB.

Technical considerations:

- Such data may be available only for certain grades (e.g., grade 3 and higher). Thus, some students who receive SES may be excluded.
- Without students’ pretest scores (e.g., those from the previous year), the gains associated with SES cannot be determined. For example, imagine that an evaluator administers a standardized math test to a student who has received SES throughout his or her 5th-grade year. Because the student scores in the 60th percentile (above the national norm), the evaluator might attribute such success to the SES provider. However, if the evaluator had tested the same student *prior to* the SES tutoring and found that he or she scored at the same level (or higher), then the evaluator’s interpretation may be quite different.
- Achievement gains based on test scores at the secondary level (e.g., using gateway exams) may be difficult to analyze because students may take tests at different times in different grades.
- If a comparison evaluation design is used, then an appropriate group of control students is needed. Students in the control group should be similar to students enrolled in SES in demographic and prior achievement variables, and they must not have received SES. (“IV. Evaluation Designs: Student Achievement” includes more details about this type of evaluation design.)
- States choosing to use state assessment results should ensure that the intended design of the assessment meshes with the purpose of the SES evaluation.

Practical considerations:

- Identifying and collecting data on control students may be expensive and time-consuming.
- In some cases, states may not have sufficient time to analyze the state assessment data. For example, if students take the exams in the spring, then states may not receive the results until mid-summer. This leaves little time to evaluate providers before the next school year.
- After following appropriate confidentiality procedures, states and/or districts may consider sharing each SES student’s pretest results with his or her chosen SES provider so that the provider understands each student’s “starting point” and can set appropriate achievement goals.

b. Supplementary individualized assessments in reading/language arts or math

States can also use student scores on other valid assessment measures (e.g., an individually administered reading inventory, such as the Woodcock–Johnson test battery or the Durrell Oral Reading test). Such tests can be administered to students enrolled in SES (and possibly control students) to assess achievement in areas in which state-mandated test data are not available (e.g., for 1st-grade students) or to obtain more in-depth or rigorous measurement of student achievement.

Technical considerations:

- As mentioned previously, without pretest scores, the gain associated with SES cannot be determined.

Practical considerations:

- Administering this option can be time-consuming and expensive for states. Specifically, the evaluator will need to obtain copies of the test, locate students receiving SES (and possibly control students), and administer and score the tests at least once.
- States using supplementary assessments may need to ensure that the timing of these tests does not conflict with state assessments or take additional time away from instruction.
- States will need to determine whether the supplementary tests will be administered at the school site (e.g., during regular classroom hours) or on the SES provider’s site and identify whose help will be needed to administer the tests.

c. Provider-developed assessments in reading/language arts or math

Some providers have developed or implemented their own assessment instruments that they use to measure student progress. Many providers have included descriptions of these assessments in their applications to the state. States can use the results from the providers’ assessments in their evaluation. This option raises critical technical and practical considerations.

Technical considerations:

- A provider’s assessment may work well for its own informational purposes but may not produce results that are statistically valid or reliable for a state’s evaluation purposes.
- If providers use different assessments, then the state will not be able to compare performance across providers.
- Assessments that are specifically tailored for a provider’s program may be more likely to show increases in student performance than an externally developed test.

Practical considerations:

- This option could save a state some time and money because it would not have to administer an assessment specifically for SES. However, sifting through results from a number of different assessments could be confusing and time-consuming.
- For more informal or “formative” evaluation purposes, states could encourage or require that providers use the same pretest and post-test to assess SES students’ progress. Because the providers would be administering those tests, the data may not be sufficiently independent of bias. However, districts and educators may find consistent formative data helpful for judging individual students’ progress, and states may find such data useful for obtaining additional, informal impressions of a provider’s impact.

2. Customer Satisfaction Measures

Parents, families, and students are an SES provider's most important customers. To collect information on customer satisfaction, states may interview or survey parents and students enrolled in SES programs. This section identifies technical and practical considerations for including parent and student perspectives in an evaluation of SES providers.

a. Parent and family perceptions

Parents and families are the provider's most important customers, and their perspectives can play an important role in developing a complete picture of a provider's impact. By law, parents choose an SES provider for their children and can move their children to another provider if they are unsatisfied with the quality of services offered. When selecting a provider for their children, parents may want to consider what other families think about each provider. To help parents make informed choices, states could publish summaries of each provider's customer satisfaction results.

Technical considerations:

- Parents and families may only be familiar with one SES provider and may not be able to compare that provider's services with the services of other providers.
- The number of students that each provider serves can vary significantly; thus, the representative sample size of parents will also vary. For example, if only a small number of students receive services from a particular provider, the impressions of their parents may not accurately characterize the provider's services on average. On the other hand, if a provider serves a large number of students, scarce time and resources may make it difficult to interview enough parents to obtain a fairly accurate picture of the different subgroups served (e.g., students who have difficulty in only math or reading, minority students, students with disabilities, ELLs).

Practical considerations:

- Identifying, contacting, and soliciting participation by the SES-involved parents may require time and effort.
- States may have to use translators to communicate with parents who do not speak English.

b. Student perceptions

Students enrolled in SES are the provider's primary customer and may be able to offer important feedback about the quality of a provider's services. Gathering and reporting student feedback on providers may help states gain a more complete picture of a provider's impact and may also help parents make informed choices.

Technical considerations:

- Students, especially younger children, may have difficulty judging the quality of services and communicating their impressions.
- This type of data can be especially time-consuming to obtain, especially if evaluators need parents' permission to survey students.

Practical considerations:

- Identifying, contacting, and soliciting participation of students enrolled in SES may require significant time and effort.
- States may have to use translators for interviews with ELL students.

3. Service Delivery Measures

States may include service delivery measures in their SES provider evaluation.³ Questions about service delivery may include: Did the provider deliver the services it promised with regard to (a) the experience and qualifications of instructors, (b) the amount of tutoring time received by students, (c) instructional strategies used, and (d) its communication with schools, districts, and parents about student progress? Did the provider meet its contractual and legal obligations with the school district? Service delivery measures also address whether a provider complies with applicable laws and contractual procedures associated with SES, including Federal, state, and local health, safety, and civil rights requirements.

To measure service delivery, states can review official records and obtain feedback from SES customers and district and school staff.

a. Records of services provided, student attendance rates, and costs

Evaluation studies cannot be meaningfully performed without a clear understanding of the exact intervention or “treatment.” State or district records will help demonstrate which services the provider actually delivered to students (e.g., the number of students served, the average number of tutoring sessions provided, and the mean length of each tutoring session). States should also consider tracking student attendance at SES sessions to accurately judge a provider’s performance. Suppose, for example, that a particular SES provider failed to show significant increase in student achievement. Examining attendance records may indicate that few students regularly attended their tutoring sessions. Therefore, the provider’s lack of success in raising achievement may merit further investigation: Is the provider making a sufficient effort to provide an SES program that engages students’ interests? Is the provider offering services at a time or in a place that is inconvenient for students or parents?

Technical considerations:

- To obtain accurate information about service delivery, states may obtain data from a variety of sources, including providers, teachers, principals, and district staff. States may choose to obtain this data during announced or unannounced site visits. During these visits, evaluators can observe providers delivering services and speak with the provider, students, families, teachers, principals, and district staff. Corroborating data from multiple sources can increase the accuracy of evaluation conclusions.

Practical considerations:

- Data may not be readily available from schools and districts. States should establish clear requirements and procedures that require districts and providers to maintain SES attendance records.
- State employee resources (or outside contractors) must be dedicated to collecting and analyzing the data.
- Onsite visits, if desired by the state, will demand significant time, labor, and resources.

³ For more details on this measure, states should check district contracts with providers and the U.S. Department of Education’s non-regulatory guidance on SES (www.ed.gov/policy/elsec/guid/suppsvcsguid.pdf). States may also consider incorporating into their evaluation elements of the Education Industry Association’s *Code of Professional Conduct for Supplemental Educational Service Providers* (www.educationindustry.org).

b. Feedback from SES customers

Measures of parent and student experiences and perceptions not only offer information about a provider's effectiveness, but they can also reveal details about a provider's service delivery.

Technical considerations:

- Parents may lack firsthand impressions or observations of the services being delivered.

Practical considerations:

- Obtaining representative samples of respondents may be difficult given mobility, availability of data, and willingness of prospective respondents to participate. Still, unless the sample is obviously non-representative, too small, or biased, the impressions provided should be useful as part of the overall evaluation.
- States may have to use translators to reach parents who do not speak English.

c. Feedback from district staff

Local school districts are a critical component of effective SES implementation. Districts contract with and pay providers and are often in closer contact with the providers than the state. As a result, districts may be able to offer crucial feedback about a provider's service delivery.

Technical considerations:

- District administrators may lack firsthand impressions or observations of tutoring services.

Practical considerations:

- Some districts may also be SES providers. Thus, a district may not be able to provide an unbiased review of its own services and may not be comfortable or able to provide an honest review of its competitors' services.

d. Feedback from school staff

Principals and classroom teachers have firsthand knowledge about a student's in-school performance. They can also offer valuable input about a provider's services. Teachers of students enrolled in SES can offer feedback about the impact of the provider's services on a student's progress and can determine whether and to what extent an SES provider works with school staff to assess the student's needs, aligns instruction with the district's instructional program, and reports back to school staff about the student's progress.

Technical considerations:

- Teachers may lack firsthand impressions or observations of providers delivering tutoring services.
- Teachers may also be SES instructors. In such cases, an instructor's dual role should be explicitly noted so that responses can be interpreted in view of possible bias.

Practical considerations:

- Identifying, contacting, and soliciting participation by the SES-involved teachers may require time and effort.
- Teachers may need to provide information about multiple providers, which may be confusing and/or time-consuming and lead to fewer surveys being completed.

Data Sources

Once states have chosen specific evaluation outcomes, they will need to design a research plan for collecting and analyzing the applicable data. Table 1 summarizes possible sources that states can use to obtain the data needed to address their respective evaluation questions.

Table 1. Summary of Data Sources by Question

Evaluation Question	Data Source
1. Does the SES program raise student achievement in reading/language arts and/or math?	<ul style="list-style-type: none"> • Pretest and post-test achievement data in reading/language arts and math.
2. Does the SES provider work with principals, teachers, and parents as needed to develop instructional plans that are geared to the needs of the students and aligned with state standards?	<ul style="list-style-type: none"> • Principal/liaison interview. • Teacher survey. • Parent survey. • Provider survey.
3. Does the SES provider communicate effectively with principals, teachers, and parents about the progress of students enrolled in SES?	<ul style="list-style-type: none"> • Principal/liaison interview. • Teacher survey. • Parent survey. • Provider survey.
4. What are teacher, student, and parent experiences with and reactions to SES interventions?	<ul style="list-style-type: none"> • Teacher survey. • Parent survey. • Student interview.
5. Does the SES provider meet the needs of ELLs and students with disabilities?	<ul style="list-style-type: none"> • District coordinator interview. • Principal/liaison interview. • Teacher survey. • Provider survey.
6. Does the SES provider deliver the services it promised in its application and/or contract with regard to (a) the experience and qualifications of instructors, (b) the amount of tutoring time received by students, (c) instructional strategies used, and (d) its communication with schools, districts, and parents about student progress?	<ul style="list-style-type: none"> • District coordinator interview. • Principal/liaison interview. • Teacher survey. • Provider survey. • Parent survey. • Onsite visit.
7. Does the SES provider meet its legal and contractual obligations?	<ul style="list-style-type: none"> • District coordinator interview. • Principal/liaison interview. • Teacher survey. • Provider survey. • Parent survey.

Technology and Database Considerations

States will need to collect a large amount of data to evaluate SES providers. To effectively and efficiently collect this data, states may consider developing and maintaining a relational database to house information that connects:

1. Achievement data and related characteristics (including demographic information) for *all* students who are eligible for SES.
2. Each student served by SES with a specific SES provider.
3. Details about the services offered by each SES provider.

The database serves as a central repository for data about SES and facilitates student-level achievement analyses. By keeping track of detailed information about providers and students, the database can be used to conduct in-depth analysis on the impact of specific provider practices such as instructional approaches, student grouping, or service duration. Analysts can also look at how a provider’s impact varies by student characteristics (e.g., grade level, race, gender, socioeconomic status, location).

Assuming that not all eligible students would participate in SES, maintaining information about “eligible but not participating” students could lay the groundwork for a possible quasi-experimental (e.g., matched control), student-level design. The information in the database could also be aggregated to provide the foundation for subsequent analyses at the school, district, and provider levels.

Data will likely need to be gathered and submitted by school or district SES coordinators. Data gathering and submission could be accomplished through a Web-based data entry mechanism or through forms that are distributed to the local points of contact and returned to the state for entry into the database. The data entry could be facilitated through standardized procedures, including a user-friendly form and predetermined submission deadlines (e.g., quarterly database updates). Here is a sample of a student information table (additional fields would likely be needed, as this is a generic example):

Sample Student Table

LEA	School Name	School ID	Student Last Name	Student First Name	Student ID	Grade Level	Receiving Services (Y/N)	Provider Name	SES Subject Area

IV. EVALUATION DESIGNS: STUDENT ACHIEVEMENT

After identifying evaluation outcomes and potential data sources, states can then begin to select the evaluation instrument(s) that will enable them to conduct an effective evaluation process. At minimum, NCLB requires that states assess student achievement to determine the provider's effectiveness. States should note that quantitative results rely greatly on the scientific rigor of the assessment's design; that is, the more rigorous the design, the more confidence can be placed in the results. Because each evaluation option requires the state to assemble comprehensive datasets and a quantitative research team dedicated to the analysis, states will also need to carefully weigh the technical merits of each design against practical considerations, such as available resources and time. For more details on maintaining data, please see the "Technology and Database Considerations" box.

This section outlines possible designs that states can use for SES evaluations. Each option has been rated from "+" (lowest scientific rigor) to "++++" (highest scientific rigor). Technical and practical considerations are also included to help states select the design that best fits their chosen evaluation outcomes and data sources.

A. Within-school matched treatment–control group with pretest and post-test

In this evaluation design, each student who receives SES is matched to a similar control student at the same school (e.g., a student who is eligible for services but does not take up the opportunity). Students are matched according to prior achievement levels, income level, ethnicity, gender, and ELL and special education status. Because it is not possible to randomly assign eligible students to receive SES, this type of quasi-experimental design is likely to be the most rigorous option available.

Rating = +++++ (high rigor)

Technical considerations:

- This type of design can show whether students who receive SES demonstrate higher achievement gains than students who do not receive SES but are similar in characteristics. By comparing two students of similar academic and eligibility backgrounds, evaluators can more confidently attribute achievement differences between groups to the intervention (i.e., SES) than to extraneous factors. A pretest score further enables statistical adjustments to be made in the analyses based on how the students performed prior to the SES intervention.
- States choosing this evaluation model should be aware that results could be skewed because of a potential selection bias between the SES-enrolled students and control students. That is, students receiving SES may possess other unmeasured characteristics that could make them more or less likely to achieve than control group students. For example, the fact that these students are enrolled in SES may be related to the possibility that they may be more motivated to do well in school or have families who are more involved in their education than students in the control group. Alternatively, teachers may be encouraging their lower-achieving students to sign up for SES.

Practical considerations:

- This design may require a significant amount of time and resources. It may be difficult to identify and match a sufficient number of students for this design. If annual state assessment results are used, then states would need to be able to match student test scores from year to year. If

supplementary individualized assessments are used, then states would need to administer the test twice (pre- and post-intervention) to two separate student groups (intervention and control).

- Although some student demographic data may be available at the state level, states may need to coordinate between schools or districts to collect the necessary data on each student.

B. Between-school matched treatment–control group with pretest and post-test

Each SES student is matched to a similar control student at a *different* school.

Rating = +++ (moderate to high rigor)

Technical considerations:

- This type of design is similar in structure to design A, except that the control student is selected from a different school than the school that the SES student attends. Design B, therefore, is less rigorous because the experiences of the SES and control students may vary more than if the students attended the same school. For example, school A (attended by the intervention student) may have adopted a new reading program this year, and school B (attended by the matched control student) may have smaller class sizes.

Practical considerations:

- This design creates a larger pool of available students to be included in the evaluation, particularly if a state is using students' results on state assessments to measure effectiveness. States using supplementary individualized assessments may need to coordinate data collection between different pairs of schools.
- Similar to design A, this design is relatively costly and time-consuming compared with less rigorous approaches (see below) that do not require a control group.

C. Regression-correlation design using pretest and post-test

Using this design, an evaluator compares a student's actual gains to predicted gains, using district-supplied data to control for student variables (e.g., income, ethnicity, gender, ELL and special education status). In other words, this design uses multiple regression analyses to generate estimates of a student's future test scores based on the student's previous test scores and other student-specific variables. If students enrolled in SES actually perform *higher* than the estimates and if students not enrolled in SES perform *at or lower* than the estimates, then researchers can reasonably infer that the SES provider shows a positive effect on student achievement. Conversely, if students enrolled in SES score *at or lower* than their predicted scores, then researchers can conclude that the provider shows no or even a negative effect.

Rating = +++ (moderate to high rigor)

Technical considerations:

- With this type of design, students enrolled in SES do not need to be individually matched to similar control students. Instead, this design simultaneously examines all available student test data, by grade, for the entire district (or a selected cohort of schools). For instance, this design can determine the extent to which SES contributes positively or negatively to student achievement by correlating achievement gains with participation in SES, while controlling for influential student

characteristics. However, such statistical controls may not be as reliable as those provided by a matched treatment-control student design (designs A and B), especially if this design is used with small sample sizes.

Practical considerations:

- Because the design works with existing datasets, states would need fewer staff resources to collect new data.
- Each grade level usually needs to be treated separately because of differences in achievement tests and student performance. As a result, the sample sizes available for a particular provider may be too small to use a regression-correlation approach.

D. Matched treatment–control group or regression using post-test only

This design is concerned with only post-test data and could be used if students’ test results are not available for the year before students enrolled in SES were (i.e., the pretest data). This situation may occur in states that do not administer assessments to primary grade students in the year before those students receive SES.

Rating = ++ (moderate rigor)

Technical considerations:

- The designs in this category have advantages and disadvantages that are similar to designs A–C but are less rigorous because they do not include pretest scores. Designs that include students’ pre-SES achievement in the analysis offer a greater amount of statistical control for reliably determining achievement gains.

Practical considerations:

- Though moderate in rigor, posttest-only designs can be low cost because they require results from only one test per student. However, states and districts need to consider the timing of these tests to ensure that students are taking the tests when they have received all or most of the SES providers’ services.

E. Treatment group-only, pretest and posttest

This design examines pre- and post-treatment gains for students who were enrolled in SES. This design does not include a control group.

Rating = ++ (low to moderate rigor)

Technical considerations:

- Because it lacks a control group, design E reduces the ability to determine the effects of SES tutoring. States should be cautious about using this design exclusively to judge provider effectiveness because the design makes it difficult to attribute gains specifically to the SES provider. For example, suppose that 70% of SES students in five schools advance to “proficient” levels in reading on the state-mandated assessment. Imagine that these five schools had also begun a new research-based reading program. Because design E does not compare similar non-SES and SES students, the evaluation may not be able to determine whether gains of SES students are

attributed to the SES provider's tutoring or to the reading program. Only by comparing the gains of students enrolled in SES with the performance of similar students *not* enrolled in SES can states infer the degree to which SES contributes to positive outcomes.

Practical considerations:

- This type of design may be less costly and require less labor, given the relative simplicity of the analyses and methodology, compared with designs A–D. However, it is much less rigorous than other design options, and states should consider using it only if designs A–D are not feasible.

F. Treatment group only, post-test only

Design F analyzes post-SES achievement for only students enrolled in SES.

Rating = + (low rigor)

Technical considerations:

- This type of design is a weaker version of design E. It's high in simplicity but low in ability to isolate specific effects of SES providers on achievement. Because this design lacks pretest (prior year) data and a control group, it would be difficult for states to determine whether student performance levels (e.g., "proficient" or "advanced") represented an actual achievement gain over time or an advantage relative to students who were not receiving SES.

Practical considerations:

- This design is simple and low cost. However, it does not allow states to accurately evaluate a provider's effectiveness and student achievement. If this design is used, states could face future costs of allowing ineffective providers to continue services, thus draining SES dollars and making little positive impact on student achievement.

V. EVALUATION DESIGNS: CUSTOMER SATISFACTION AND SERVICE DELIVERY

Table 2 summarizes possible sampling strategies and data collection tools⁴ that states can use to assess an SES provider's customer satisfaction and service delivery outcomes. In general, the table suggests that sample sizes of 5 to 15 respondents are sufficient for interviews that measure "impressions" of programs from SES providers. Some states may use a percentage of students served to set their sample sizes. States using percentages may need to conduct a large number of interviews with providers that serve a large number of students. Using percentages can also undesirably reduce the sample size for providers that serve a small number of students.

⁴ Scannable surveys have been developed by the Center for Research in Educational Policy and are available for review. For more information, go to <http://crep.memphis.edu>.

Table 2. Sampling Strategies and Data Collection Tools to Assess Customer Satisfaction and Service Delivery

Respondent Group	Sampling Strategy	Data Collection Tool
District SES Coordinators	One coordinator per district	Phone interview with district coordinators
Principals or school representatives	At least one principal or school representative per district, to the extent that the respondents are familiar with each provider. States can stratify the sample based on the number of SES schools.	Phone interview with or survey of principals or school representatives
Teachers	Eight to 15 teachers per provider, depending on the number of students served	Survey of teachers
Parents	Five to 10 parents per provider, depending on the number of students served	Phone interview with or survey of parents
Providers	One representative from each provider	Survey of providers
Students	Five to 10 students per provider, depending on the number of students served	Interview with students

VI. OVERALL PROVIDER EVALUATION RUBRIC

A state’s overall evaluation of an SES provider should link all three dimensions of performance:

1. **Effectiveness.** Measured by changes in student achievement in reading/language arts or math.
2. **Customer satisfaction.** Measured through surveys and interviews with parents and students who receive SES.
3. **Service delivery.** Measured through surveys and interviews with principals, teachers, and school district staff to determine compliance with applicable laws and contractual procedures associated with the delivery of SES.

States will need to synthesize data from multiple sources perform a “summative” evaluation of each provider. States can use the results from the summative evaluation to determine which providers can maintain their “state-approved” status. Figure 1 provides a graphic representation of a summative evaluation.⁵

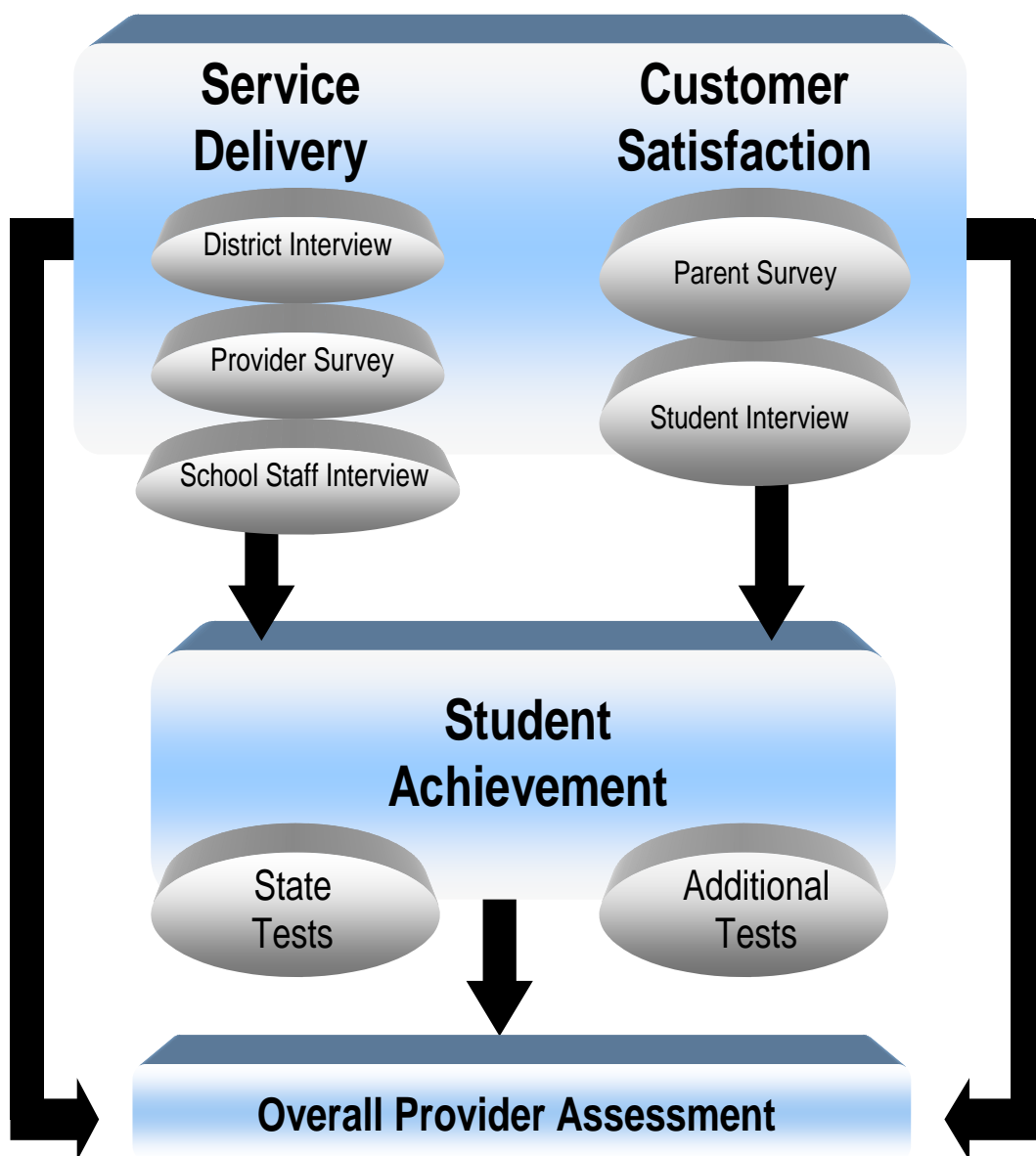
To facilitate this process, states may choose to develop a culminating tool, or an “overall provider evaluation rubric,” to synthesize the multiple data sources that reflect providers’ activities and outcomes. This tool can help states clarify how well each provider meets expected outcomes and standards. The rubric can be organized by the major assessment categories (e.g., student achievement gains, effective communications) that are included in a state’s SES evaluation plan; and each category should be defined. For each category, a scoring rubric—consisting of four levels of attainment (e.g., above standards, acceptable, marginally acceptable, and below standards) and accompanying written definitions—can be

⁵ This model is also used in the Center for Research in Educational Policy’s comprehensive SES Evaluation/Monitoring Plan; see <http://crep.memphis.edu>.

constructed. For example, an “above standards” level for student achievement gains and for effective communications may appear as follows:

- **Student achievement gains.** SES students’ average gain shows an effect size of +0.XX⁶ or greater, relative to the average gain for control students or as indicated in appropriate norms.
- **Effective communications.** Based on interviews, *all* of the district’s coordinators, school principals, liaisons, teachers, and parents rate the provider positively in communicating SES activities, student needs, and progress. The provider also produces *concrete* evidence supporting usage of effective communications with these groups.

Figure 1. Components of a Comprehensive SES Evaluation Plan



⁶ States will need to determine the effect size.

VII. COMMUNICATING RESULTS

States will need to communicate the results of provider evaluations to providers, families, schools, districts, and the general public. A state should consider whether and how it will report results in a user-friendly way. States may provide more detailed information to providers to help them understand strengths and weaknesses of their SES program and make needed improvements. States may also choose to post summaries of each provider's evaluation results on the state Web site so that districts and families can access the information and use it to select a provider. To help ensure that information is provided in a central, accessible place, states can incorporate these results into their publicly available lists of approved SES providers. States should also consider whether they will share evaluation data with researchers who are interested in assessing the overall impact of SES.

VIII. CONCLUSION

The evaluation of SES providers is a new and emerging endeavor, and evaluation models are still evolving. As is true in virtually all educational domains, no one evaluation approach is likely to fit all conditions and interests. Each state has unique needs, priorities, resources, and procedures that it can use to implement SES. Therefore, this Issue Brief presents a range of recommended outcome measures and research designs for states to consider. The options presented focus strongly on the critical goal of assessing the effects of SES on student achievement.

As the descriptions of the different evaluation options convey, states may face a trade-off between practical concerns (cost and time) and rigor (the reliability and accuracy of findings). As a general approach, each state should begin its SES evaluation planning process by:

- Identifying, through review of this issue Issue Brief and discussions with key stakeholders, the specific questions that its SES evaluation needs to answer.
- Identifying the resources that can be allocated reasonably to support further evaluation planning, data collection, analysis, reporting, and dissemination.

With a clear understanding of its research needs and practical resources, a state can work through the hierarchy of evaluation designs presented here and select the design that allows the highest level of rigor. In some cases, states may adopt a hybrid approach in which (a) a less rigorous but more practical design (e.g., treatment group only, pretest and post-test) is broadly applied across the state and (b) a highly rigorous design (e.g., within-school matched treatment–control group, pretest and posttest) is used in selected contexts and then expanded each year. Evaluations may also include an analysis of providers' outcomes relative to the cost of their services. States may opt to engage third-party evaluation experts in helping to plan and conduct these evaluations. Such experts can lend experience, knowledge, and credibility to the evaluation, if the evaluator is not biased toward a particular SES provider, district, or school. The What Works Clearinghouse's *Registry of Outcome Evaluators* is one source that states can use to find an evaluation expert. To search the registry, go to www.whatworks.ed.gov.

Given the importance and complexity of evaluating SES providers, it makes little sense for each state to reinvent the wheel. Over time, SES evaluation experiences nationally will yield valuable information and improved data collection methodologies and tools. By exchanging tools and lessons learned with other states, officials can save considerable time and resources, but most importantly, they can improve the validity, practicality, and usefulness of their own SES evaluations.

IX. ADDITIONAL RESOURCES

American Educational Research Association (AERA)

www.aera.net

AERA is an international professional association of researchers that is focused on advancing educational research and its practical applications.

American Evaluation Association (AEA)

www.eval.org

AEA is an international professional association of evaluators. It is focused on improving evaluation practices and methods that are used to assess the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness.

Center for Research in Educational Policy (CREP)

<http://crep.memphis.edu>

Based at the University of Memphis, CREP conducts research on educational policies and practices in pre-K–12 public schools and provides evaluation instruments for educational practitioners. CREP's Dr. Steven Ross co-authored this Issue Brief with the SESQ Center.

Council of Chief State School Officers (CCSSO)

www.ccsso.org

CCSSO works with chief state school officials and their agencies to strengthen and support the American education system. Its Web site offers guidance and resources on afterschool programs and SES.

Education Leaders Council (ELC)

www.educationleaders.org

ELC is a network of educational leaders that work to improve academic achievement for all students. ELC offers guidance to states on how to create and manage SES provider evaluation systems.

National Reporting System (NRS) for Adult Education

www.nrsweb.org

NRS is an outcome-based reporting system for the state-administered, federally funded adult education programs. The Web site offers guidance on collecting, analyzing, and reporting the impact of adult education programs.

North Central Regional Educational Laboratory: More on Evaluation Standards

www.ncrel.org/tandl/eval6.htm

This North Central Regional Educational Laboratory tool outlines four key evaluation principles to consider when creating program evaluations.

Office of Innovation and Improvement

www.ed.gov/about/offices/list/oii/index.html?src=oc

This office of the U.S. Department of Education coordinates public school choice and supplemental educational services programs along with the Office of Elementary and Secondary Education (www.ed.gov/about/offices/list/oese/index.html).

Practical Assessment, Research & Evaluation

<http://pareonline.net/>

This online journal provides education professionals with access to peer-reviewed articles on assessment, research, evaluation, and teaching practice, especially at the district level.

Regional Educational Laboratories (REL)

www.relnetwork.org

REL is a network of 10 educational laboratories that serve specific geographic regions of the United States. The laboratories provide information, research, and tools to help those involved in educational improvement at the local, state, and regional levels.

Supplemental Educational Services Non-Regulatory Guidance

www.ed.gov/policy/elsec/guid/suppsvcsguid.pdf

This resource serves as the U.S. Department of Education's guidance on SES. It provides complete information for states and districts to implement provisions of SES.

The Evaluation Exchange

www.gse.harvard.edu/hfrp/eval.html

This periodical from the Harvard Family Research Project highlights innovative methods and approaches to evaluation, emerging trends in evaluation practice, and practical applications of evaluation theory. Subscriptions are free, and it is published three to four times per year.

User-Friendly Handbook for Project Evaluation

www.nsf.gov/pubs/2002/nsf02057/start.htm

This handbook from the National Science Foundation explains how to create and implement evaluations for educational programs. It offers information on quantitative and qualitative evaluation methods and guidance on locating external evaluators.

What Works Clearinghouse (WWC)

www.whatworks.ed.gov

WWC collects, screens, and identifies studies of the effectiveness of educational interventions. Online visitors can search the Clearinghouse's *Registry of Outcome Evaluators* to find an independent evaluator that may be able to assist with SES evaluations.

ACKNOWLEDGEMENTS

This document was collaboratively developed by Dr. Steve Ross from the Center for Research in Educational Policy, at the University of Memphis, and the Supplemental Educational Services Quality (SESQ) Center. We would like to express our appreciation to those who helped contribute to this Issue Brief, including staff of the American Institutes for Research and the Center for Research in Educational Policy; Janelle Cousino of Fowler-Hoffman, LLC; Ayeola Fortune of the Council of Chief State School Officers; Stephanie Gerber of the Ohio Department of Education; Kay Kammel of the Florida Department of Education; Steve Pines of the Education Industry Association; and Ted Rebarber and Gary Huggins of the Education Leaders Council. We also extend thanks to the SESQ Coordinating Group for their valuable input. The development of this Issue Brief was supported through a grant from the U.S. Department of Education. The views expressed herein do not necessarily represent the positions or policies of the U.S. Department of Education. No official endorsement by the U.S. Department of Education is intended or should be inferred.



About the SESQ Center

The Supplemental Educational Services Quality (SESQ) Center helps low-income families take advantage of a new opportunity, provided by the *No Child Left Behind Act*, to get their children free tutoring and the extra academic help they may need. The SESQ Center was established through a grant to the American Institutes for Research from the Office of Innovation and Improvement of the U.S. Department of Education.

To meet the challenges of implementing supplemental educational services, the SESQ Center works to:

- Increase the number of eligible children receiving free tutoring.
- Expand the number and range of high-quality tutoring services offered.
- Improve district, state, and national coordination of the program.

The SESQ Center provides technical assistance at local and state levels through a network of demonstration districts and states. At the national level, the SESQ Center created a network of organizations to coordinate and improve research and assistance on supplemental services. The SESQ Center Web site, www.tutorsforkids.org, turns “lessons learned” into tools and resources on effective supplemental services implementation that can be replicated by other districts and states.

Web site: www.tutorsforkids.org
Toll-free phone: (866) 544-8686