

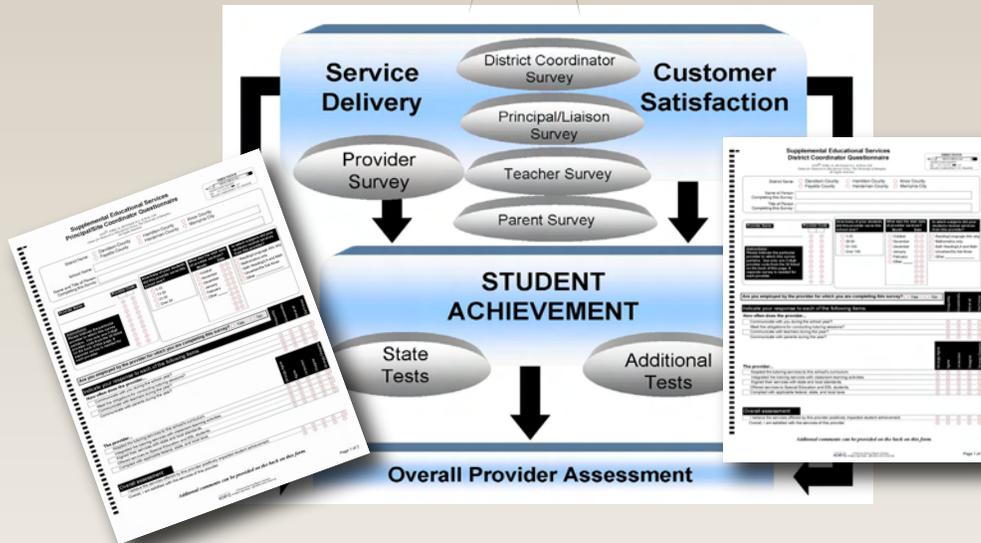
EVALUATING SUPPLEMENTAL EDUCATIONAL SERVICE PROVIDERS: SUGGESTED STRATEGIES FOR STATES (2ND EDITION)

A Guidebook Prepared for the

CENTER ON
INNOVATION &
IMPROVEMENT

Twin paths to better schools

INFORMATION TOOLS TRAINING



Information Tools Training

Positive results for students will come from changes in the knowledge, skill, and behavior of their teachers and parents. State policies and programs must provide the opportunity, support, incentive, and expectation for adults close to the lives of children to make wise decisions.

The Center on Innovation & Improvement helps regional comprehensive centers in their work with states to provide districts, schools, and families with the opportunity, information, and skills to make wise decisions on behalf of students.

*A national content center supported by the
U. S. Department of Education's Office of Elementary and Secondary Education.
Award #S283B050057*

*The opinions expressed herein do not necessarily reflect the position of the supporting agencies,
and no official endorsement should be inferred.*

TABLE OF CONTENTS

I.	About this Guidebook	3
II.	Guiding Questions for Evaluating Supplemental Educational Service Providers	5
III.	Determining Evaluation Measures.....	6
IV.	Evaluation Designs: Student Achievement.....	15
V.	Evaluation Designs: Customer Satisfaction and Service Delivery.....	18
VI.	Provider Evaluation Rubric and Decision Tree	19
VII.	Communicating Evaluation Plan Results.....	21
VIII.	Conclusion	22
IX.	Additional Resources.....	22
X.	Acknowledgments and Disclaimers.....	24
Appendix A. Sample Instruments/Tools		28
Appendix B. Rubric of Overall Provider Effectiveness.....		40
Appendix C. Decision Tree for SES Providers		41
Appendix D. Descriptions of Provider Status Categories		42

I. ABOUT THIS GUIDEBOOK

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

The first edition of this Guidebook was published in spring 2005 by the Supplemental Educational Services Quality Center, which was operated by the American Institutes for Research[®] (AIR[®]) from 2003–2005. Following the publication of the first edition of the Guidebook, much experience was gained and lessons were learned through the initial evaluation studies that the present authors and other groups conducted. At the same time, stakeholders in the evaluation process (e.g., policy makers, regulators, educators, providers, and evaluators) have gained an increased understanding of expectations, challenges, and potentially useful evaluation models. This revised edition of the Guidebook reflects these advances in the field. For example, we now propose, even more strongly than we did previously, that evaluation designs should supplement the mandatory examination of student achievement outcomes with performance dimensions addressing customer satisfaction and service delivery, including provider efforts to promote regular student attendance. Like the first edition, this Guidebook explains these performance domains and describes possible evaluation measures and designs that states can use to assess each domain. This version goes one step further by proposing strategies for (a) “front-end” communications with providers about evaluation requirements and expectations, and (b) end-of-year decisions about provider status (continuation or removal) based on the evaluation results.

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 (NCLB) Act. 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 2nd year of “school improvement” status). A variety of organizations are eligible to provide SES after receiving formal approval from the state, such as:

- * 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 improvement¹

Roles of States and Districts

NCLB and the U.S. Department of Education (ED) 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 SES provider application and criteria.

¹The U.S. Department of Education (ED) recently granted waivers to a select number of districts identified for improvement to allow those districts to provide SES. For more information about this pilot, go to: <http://www.ed.gov/news/pressreleases/2006/07/0762006a.html>.

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.² 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 SES provider application. 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 SEAs may also request that districts help monitor SES providers.

Districts are charged with determining which students in Title I schools identified as in need of improvement are eligible for SES.³ 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 a SES provider. If families’ demands for SES exceed the available funds, districts must then give priority to the lowest-achieving eligible students. Once families choose a provider, then the district contracts with the provider and pays for its services. When establishing contracts, districts and providers should work together to clearly indicate requirements and expectations of the SES evaluation, particularly with regard to how “effectiveness in increasing student achievement” will be determined.

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 when data are needed regarding: (a) student participation in SES; (b) student attendance; and (c) experiences and satisfaction with provider services by district SES liaisons/coordinators, parents, and teachers and principals from participating schools. In such instances, states should ensure that the district’s role is appropriate and unbiased, because school districts may also be SES providers. States and districts can also collaborate to ensure that district contracts with providers establish procedures that states need to gather data and evaluate providers. Additionally, many of the evaluation options discussed in this Guidebook require consistent, statewide data gathering and reporting mechanisms. To ensure consistency and availability of required data for monitoring and evaluating SES, states are strongly encouraged to implement a statewide data collection and compilation system. (For more guidance on data collection options, consult the Technology and Database Considerations box on page 15.)

²States are encouraged to monitor providers at least annually.

³If 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. Currently, ED is exploring the viability, through waivers, of offering states flexibility in choosing the order of implementing the transfer and SES options to qualifying students. Additional information about NCLB’s choice options is available from ED’s Web site: <http://www.ed.gov/parents/schools/choice/definitions.html>.

II. GUIDING QUESTIONS FOR EVALUATING SUPPLEMENTAL EDUCATIONAL SERVICE 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 are encouraged to 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 of students who received SES satisfied?
3. **Service delivery and compliance.** 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 providers are expected to demonstrate effectiveness across all of the types of students served, including English language learners (ELLs) and students with disabilities and a variety of learning needs. However, as we will discuss in the following sections, a realistic constraint to evaluating such effectiveness meaningfully and validly is that individual providers and grade levels of interest may have only small sample sizes for test scores necessary for analysis. Accordingly, the orientation that we recommend throughout the Guidebook is a pragmatic one based on the following premises:

- * Evaluation is useful for decision making only when based on valid data.
- * SES evaluations will be restricted in any given year by the number of participating students for which appropriate (valid) achievement data are available.
- * Each year that achievement data are collected, the reliability of information about providers overall and for different student subgroups increases.
- * Over time, the ability to make confident and scientifically valid decisions about provider effectiveness will increase.

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 related to:

- * Tutors' experience and qualifications
- * The amount of tutoring time students receive
- * The teaching strategies used
- * Instructional grouping and student–instructor ratios
- * Communication with teachers and parents
- * Promised transportation of students to and from tutoring

In our prototype evaluation model (see section VI), we include such assessments in the Compliance category. Depending on the severity of compliance violations, providers may be downgraded from the highest status rating (e.g., Full Standing) to a lower, probationary status, or in extreme cases, removal.

Basic Components of a Comprehensive Evaluation

- * Communicate evaluation plan to all stakeholders, including providers. See Section VII. Communicating Evaluation Plan and Results.
- * Collect customer satisfaction, compliance, and achievement data during the school year. See Section III. Determining Evaluation Measures.
- * Synthesize data from multiple sources to rate the provider’s performance on key categories (e.g., Student Achievement, Communications, Instructional Plans, etc.). See Section VI. Provider Evaluation Rubric and Decision Tree.
- * Obtain interrater corroboration on ratings, particularly those indicative of unsatisfactory performance. See Section VI. Provider Evaluation Rubric and Decision Tree.
- * Use ratings in different categories and a formal “decision” process to judge provide overall status for the next year. See Section VI. Provider Evaluation Rubric and Decision Tree.
- * Communicate results in appropriate manner to the state, districts, providers, and the public. See Section VII. Communicating Evaluation Plan and Results.

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.

A. 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.

1. 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., Grades 3 to 8). 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.

Why Are Pretest Scores Important?

Example A: A third-grade student who received SES tutoring during the school year scores in the 60th percentile (above the national norm) on the spring state testing. An initial impression is to view the SES provider as having contributed positively to this relatively high performance. Although this might be the case, the child was not tested as a second-grade student the prior year. Perhaps, if she had been tested in second grade, she would have performed comparably or even relatively higher (e.g., the 65th percentile). In that case, the evaluator's interpretation would be quite different.

Example B: Using a treatment control group evaluation design, the evaluator compares 125 second-grade students who received SES from Totally Terrific Tutoring, Inc. (T3) to 328 SES-eligible second-grade students from the same district who did not elect to receive tutoring. The comparison group significantly outperforms the SES group on the state-mandated achievement test. T3 is very disappointed with this finding. However, had the second grade students been pretested (prior to SES tutoring), the evaluator would have found that the SES students were much more disadvantaged and lower performing than the control group. Was the T3 tutoring actually ineffective?

- * Achievement gains based on test scores at the secondary level (e.g., using benchmark, end-of-course, or gateway exams) may be difficult to analyze because students may take tests at different times, in different grades, and multiple times.
- * Although a comparison evaluation design is usually desirable, 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. Section IV. Evaluation Designs: Student Achievement includes more details about this type of evaluation design. If data from a suitably large sample of SES and non-SES are available, it may be feasible to create a comparison group by adjusting statistically for individual and group differences in ability and background in the analyses of achievement scores (see below).

Practical considerations:

- * Identifying and collecting data on control students may be expensive and time-consuming. As discussed in section IV, a more practical (but probably less powerful) evaluation design is a multiple regression-type study that collects data on all students (both SES and non-SES) in a given subpopulation, such as those enrolled in Title I schools in the districts concerned. The analysis determines whether SES participation is associated with higher-than-expected achievement scores after controlling for student demographics (e.g., gender and ethnicity) and prior achievement scores.
- * In most cases, states may not have sufficient time to analyze the state assessment data before providers are approved to offer services for the next school year. For example, if students take the exams in the spring, then states may not receive the individual student score results until mid-summer. This leaves little time to analyze the data and evaluate providers before school begins.
- * To permit rigorous analyses, students' actual scores (e.g., number correct, scale score, or standard score) on the state assessments are needed. Merely knowing their proficiency levels,

such as “advanced” or “below basic,” usually will not provide the degree of precision needed to detect differential growth for SES versus non-SES students.

2. 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 comparison students) to assess achievement in areas in which state-mandated test data are not available (e.g., for first-grade students), or to obtain more in-depth or rigorous measurement of student achievement.

Technical considerations:

- * As mentioned previously, without pretest scores, the gain or benefit associated with SES cannot be determined.
- * Without comparison students, there will be no basis for judging whether SES students are scoring higher than peers who did not receive the tutoring services. For example, high scores from SES students could mostly be a reflection of the state’s improved reading standards and curriculum, or some other systemic reform.
- * Unless these assessments are administered by trained independent testers, their validity may be suspect. Specifically, classroom teachers may lack adequate training or be unable to follow the protocols (e.g., completion time, room conditions, etc.) required for administering tests accurately and consistently across students. SES tutors, obviously, would be put in an awkward position if they were responsible for administering the very tests from which they and their programs would be judged.

Practical considerations:

- * Administering supplementary tests can be very time-consuming and expensive for states. Specifically, the evaluator will need to obtain copies of the test, locate students receiving SES (and presumably comparison 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. Note that if the tests are administered at the provider’s site, equivalent testing for comparison students would be precluded.

3. Provider-Developed Assessments in Reading/Language Arts or Math

Some providers have developed or implemented their own assessment instruments to measure student progress. Many providers have included descriptions of these assessments in their applications to the state. Use of such diagnostic and formative tests is clearly a positive program component and should be strongly encouraged. However, the question of whether the results from the providers’ assessments should be a formal component of the states’ evaluation of provider effectiveness 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 valid or suitable for a state’s evaluation purposes. That is, high performance on a provider’s test may not translate necessarily into proficient performance on the state’s test.
- * 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 would an externally developed test.
- * As noted above, the objectivity and validity of scores may be compromised when the providers themselves are asked to administer and score tests that will be used to judge their effectiveness. For many providers, what had been viewed as highly valuable diagnostic and formative tests would now become “high stakes” assessments with significant implications for their status and reputation.

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 providers to use the same assessment instruments to measure SES students’ progress. Because the providers would be administering those tests, the data may not be sufficiently independent of bias, as discussed previously. However, districts and educators may find consistent, formative data helpful for judging individual student progress, and states may find such data useful for obtaining additional, informal impressions of a provider’s impact.

B. Customer Satisfaction Measures

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

1. Parent and Family Perceptions

Perspectives from parents or caretakers can play an important role in developing a complete picture of a provider’s impact. By law, parents choose a SES provider for their children and can move their children to another provider if they are dissatisfied with the quality of services offered. When selecting a provider for their children, parents may want to consider what other families think about that provider. To help parents make informed choices, states could publish summaries of each provider’s customer satisfaction results (see section VII). A sample survey that CREP/EI has used in several state evaluations is provided in Appendix A. Sample Instruments/Tools.

Technical considerations:

- * Parents and families may be familiar with only one SES provider, and may they not be able to compare that provider’s services with the services of other providers. For that matter, except for the rare situations where SES tutoring occurs in the home, parents may have little or no impression of the quality or nature of the services provided.
- * 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.
- * Parents who participate in surveys and interviews may not represent the population of parents whose children the provider serves (see next section).

Practical considerations:

- * Identifying, contacting, and soliciting participation by the SES-involved parents may require extensive time and effort.
- * In the interest of protecting the confidentiality of SES participants, conducting interviews

would generally be precluded. Even if approval was obtained, interviews would be logistically challenging for the evaluators. Many parents do not like strangers contacting them by phone to ask questions about their child. Probably, even fewer would want to take the time to attend an in-person interview or have an interviewer show up at their home.

- * Paper-and-pencil (not electronic) questionnaires seem the most practical way of obtaining parent reactions. The problem with electronic surveys is that many parents (in this predominantly low-socioeconomic status population) may not have computers or be computer literate. Due to confidentiality issues, the evaluators may not be approved to mail or distribute the surveys to the parents. Therefore, we have found that the most viable procedure is mailing the surveys to a school liaison who then mails or sends them via the students' "homework packets" to the parents. The parents, in turn, return the surveys to the school liaison who then mails them to the evaluators.
- * States may have to use translators to communicate with parents who do not speak English.

2. 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. However, as conveyed below, we have serious reservations about the viability of surveying/interviewing students from both a logistical and validity (accuracy) standpoint.

Technical considerations:

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

Practical considerations:

- * Identifying, contacting, and soliciting participation of students enrolled in SES may require significant time and effort.
- * If confidentiality requirements prohibit the evaluators from knowing the identity of SES students, questionnaires would need to be administered for the evaluator by school or district personnel. This would significantly increase the logistical demands for administering questionnaires while precluding interviews. It seems almost certain that sample sizes will be small and nonrepresentative of the population receiving the individual provider's services.
- * States may have to use translators for interviews or questionnaires with ELL students.

C. Service Delivery and Compliance Measures

States may include service delivery measures in their SES provider evaluation.⁴ 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.

⁴For more details on this measure, states should check district contracts with providers and ED's nonregulatory 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).

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

1. Records of Services Provided, Student Attendance Rates, and Costs

Evaluation studies cannot be performed meaningfully 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 track student attendance at SES sessions to accurately judge a provider’s performance. Suppose, for example, that a particular SES provider failed to show significant increases 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 implement a 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?

Tracking student attendance and total hours of tutoring received is also important for increasing the rigor of analyses of provider effectiveness in raising student achievement (see Section IV. Evaluation Designs: Student Achievement). That is, including in such analyses students who have received only a few hours of tutoring would bias the results. But what is “minimal” or “adequate” tutoring? States need to define such levels.

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 data during announced or unannounced site visits. During these visits, which may be part of routine monitoring by state staff, the observers can view 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 procedures that require districts and providers to maintain SES attendance records. Using an automated, centrally controlled (by states), Web-based data recording system is strongly recommended. Several commercial vendors offer services and software for these purposes.
- * State employees (or outside contractors) must be dedicated to collecting and analyzing data.
- * Onsite visits, if desired by the state, will demand significant time, labor, and resources.

2. 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:

- * As previously discussed, 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 nonrepresentative, 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.

3. 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. Online surveys, similar to samples in Appendix A. Sample Instruments/Tools, are used for SES coordinators in each district with schools required to offer SES.

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.

4. 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. In particular, teachers of students enrolled in SES can offer feedback about the impact of the provider's services on a student's progress, and they can determine whether and to what extent a 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. A sample survey instrument (CREP, 2006) is shown in Appendix A. Sample Instruments/Tools.

Technical considerations:

- * Teachers may lack firsthand impressions or observations of providers' service delivery.
- * Teachers may also be SES instructors. In such cases, a teacher'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 for the school. Typically, the evaluator will not know the identities of the teachers involved, and will mail the surveys to the school for distribution internally.
- * Teachers may need to provide information about multiple providers, which may be confusing and time-consuming, leading 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 posttest 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 survey • District coordinator survey • 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 survey • District coordinator survey • 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 survey
5. Does the SES provider meet the needs of ELL students and students with disabilities?	<ul style="list-style-type: none"> • District coordinator survey • Principal/liaison survey • 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 students receive, (c) instructional strategies used, and (d) its communication with schools, districts, and parents about student progress?	<ul style="list-style-type: none"> • District coordinator survey • Principal/liaison survey • Teacher survey • Provider survey • Parent survey • Monitoring documentation
7. Does the SES provider meet its legal and contractual obligations?	<ul style="list-style-type: none"> • District coordinator survey • Principal/liaison survey • Teacher survey • Provider survey • Parent survey • Monitoring documentation

IV. EVALUATION DESIGNS: STUDENT ACHIEVEMENT

After identifying evaluation outcomes and potential data sources, states can 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 in the tutored subjects (reading/language arts and/or mathematics). 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 on page 15.

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. However, in describing different options, we want to strongly emphasize the importance of states making every effort to implement a design that is rigorous (“+++” or “++++”). A basic tenet of educational research is that weak evidence can be more harmful than having no evidence at all. Think of the consequences to children, schools, and the provider industry of erroneously identifying individual providers as effective or ineffective. Newspapers and other media are often quick to provide exaggerated characterizations of outcomes without explaining the nuances, as reflected in a headline such as “SES Tutoring Fails to Raise Student Achievement.” Quite possibly, the evaluator simply compared SES and non-SES students on the spring state assessment, without adjusting for the higher poverty and lower prior achievement of the students enrolled in SES. A more rigorous evaluation might have demonstrated that SES tutoring was highly effective at fostering year-to-year gain.

A. Benchmark Comparison

In this evaluation design, the aggregate percentage of students attaining “proficiency” on the state assessment is computed for each provider and descriptively compared to established benchmarks set by the state. Confidence intervals can be constructed around each provider estimate to determine statistical reliability. In general, this design is not sufficiently rigorous for making judgments of provider effectiveness that would result in major status decisions (e.g., removal of a provider from a state-approved listing). However, where resources or time are limited, it could provide informative “suggestive” impressions about tutoring outcomes. A more rigorous design might then be implemented the following year.

Example: Provider A has provided math tutoring to 254 students who have test scores in mathematics on the state assessment. Of these students, 82% score at the “proficient” level or higher, which exceeds the state benchmark for the year of 76%. Computation of a confidence interval around the estimate supports the conclusion of a statistically significant advantage for the Provider A students.

Rating = + (low rigor)

Advantages:

- * Relatively inexpensive from a data processing and analysis perspective.
- * Easily understood by practitioners and the public.

- * The same clear standard, directly linked to NCLB accountability, is applied to all providers.

Disadvantages:

- * Confounding of provider and school effects. For example, Provider A may have a large proportion of its students clustered in a small number of schools that are adopting new research-based curricula.
- * Does not capture student growth along a score continuum (e.g., 0 to 100). That is, performance categories such as “proficient” or “not proficient” can include a wide range of scores, therefore sacrificing precision of measurement.
- * Does not take into account variability in outcomes that may be strongly related to individual student characteristics. For example, Provider A might be serving a higher risk student population (lower achieving, ELL, and special needs students, etc.) than other providers.
- * May create adverse incentives; that is, providers avoid serving students who are most at-risk. To the extent that “proficiency” is established as the criterion for success, moving students from very low “below basic” levels to higher (but still below proficient) levels will not be considered.

Rating = ++ (low to moderate rigor)

Strategies for improving this design would include:

- * Including all performance categories (e.g., below basic, basic, proficient, or advanced) rather than only the percentage meeting benchmark
- * Determining SES students’ improvement in performance relative to the prior year and comparing the improvement (or gain) score to that for the state overall
- * Comparing yearly performance or gain of SES students to a control sample of similar non-SES students

B. Multiple Linear Regression Design (Using Pretest and Posttest)

Using a statewide database, a Multiple Linear Regression (MLR) model is estimated in which the current year’s achievement score is predicted by prior achievement, free or reduced-price lunch status, and other pertinent predictor variables. A “fixed effect” (effectiveness score) for each provider can be estimated and subjected to an inferential statistical test. This test would determine whether the effect is statistically reliable rather than likely to occur by chance.

More simply, this design uses correlation-type analyses to generate estimates of a student’s future test scores based on his/her previous test scores and other student demographic 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, researchers can conclude that the provider shows no or even a negative effect.

Example: There are 45 fourth-grade students who receive reading tutoring from Provider B. Based on the students’ third-grade reading scores, gender, and ethnicity, the MLR model estimates how each is likely to perform on the fourth-grade assessment. Students’ actual scores are then compared to their predicted scores. Overall findings indicated that the average student scored 10 points higher than predicted! The 10 points reflect a one-half standard deviation advantage, which translates into an effect size of +0.50 (very strong). Further, the effect size is determined to be statistically reliable, with less than a 1% probability of occurring by chance. The provider therefore is evaluated as “effective” in raising student achievement.

Rating = +++ (moderate rigor)

Advantages:

- * Although more expensive than the Benchmark Comparison, the MLR approach is still relatively economical with regard to time and cost. That is, students enrolled in SES do not need to be individually matched to similar control students (see design C, Matched Samples Design).
- * By taking student characteristics into account, a more precise determination of the provider's unique contribution ("effects") can be determined. Note that because prior test scores are the strongest correlate of subsequent achievement, the absence of pretest data will reduce the rigor of the MLR (and other design options) considerably. In fact, we recommend that definitive judgments about provider performance not be made on the basis of posttest-only designs.
- * The effect size estimates (see above example) can be combined across grades and years via meta-analytic procedures to increase power to detect small provider effects over time.

Disadvantages:

- * Confounding of provider and school effects. For example, Provider A may have a large proportion of its students clustered in a small number of schools that are adopting new research-based curricula.
- * Less easily understood by practitioners and the public than the Benchmark Comparison.
- * The achievement estimates may be less stable and efficient than those derived from more sophisticated approaches (see design C, Matched Samples Design).

C. Matched Samples Design

There are a variety of approaches to creating matched samples. The simplest approach involves matching each SES student with a similar student attending the same school. Provider effects are then computed relative to this comparison sample using inferential statistical analyses.

Example: Provider C has tutored 78 students in reading across the state. For each of the 78 students, the evaluator identifies a matched control student counterpart who (a) attends the same school or, if not possible, a similar school in the same district; (b) is the same in gender, poverty (free or reduced-price lunch status), and ethnicity; and (c) scored identically or very similarly in reading on last year's state assessment. Reading scores on this year's state test are then compared for the matched-pair SES and control groups.

Rating = ++++ (high moderate to strong rigor)

Note that greater precision and efficiency can be achieved through use of multiple matches per SES student. This procedure might involve matching SES students to highly similar students within their school and to those attending other similar schools. Thus, for example, if SES student # 306 matched well to five non-SES students at her school and two neighboring schools, it would increase statistical power and precision to use the multiple matches rather than only one. Another means of increasing precision is using more sophisticated statistical procedures that match students via "propensity score weights."

Advantages:

- * Provides some control over school effects (where there are within-school matches)
- * Conceptually easy to understand and has strong "face validity"

- * Permits computation of traditional effect size estimates
- * Probably the most rigorous approach available

Disadvantages:

- * Substantially more time-consuming than the Benchmark Comparison or MLR approaches from a data analysis perspective.
- * In some cases, there may be too few appropriate within-school matches, thus reducing ability to control for school effects. This problem can be ameliorated by using more complex matching strategies, but the straightforward appeal of simple within-school matching would be compromised.

V. EVALUATION DESIGNS: CUSTOMER SATISFACTION AND SERVICE DELIVERY

Table 2 summarizes possible sampling strategies and data collection tools⁵ that states can use to assess a SES provider's customer satisfaction and service delivery outcomes. In general, Table 2 suggests that surveys or interviews from at least five parent and teacher respondents per provider should be obtained when possible. A larger and more representative sample will result in greater confidence in the validity of results. In many instances, evaluators will need to decide subjectively and defend the level of confidence that can be attached to the data. In the case of parent surveys in particular, obtaining "representative" samples will be very challenging or not even possible. Specifically, respondents are likely to be disproportionately those who (a) are more involved with the children's education, (b) are more literate or educated, or (c) have particular positive or negative feelings or personal agendas regarding the tutoring services.

How many respondents are enough? The number of students served may vary considerably across providers within a state. One option is to establish target sample sizes based on a percentage of students served. However, the result may be unreasonably large or small numbers for particular providers. We prefer a pragmatic approach that strives to include as many survey respondents as feasible in order to increase sample size. Below are two scenarios in which results from small samples might be interpreted differently by the evaluator.

Example 1: Provider A serves 32 children in a school district. Of the corresponding 32 surveys sent home to parents, only 8 are returned. However, all 8 show consistently positive ratings, while including open-ended comments noting specific beneficial activities of the provider (e.g., calling the parent, relating well to the student, and increasing motivation and learning). These impressions corroborate those from other sources. Although sample size is small, the evaluator feels confident in interpreting these results as highly supportive.

Example 2: Same as example 1, but the eight returned surveys reveal highly varied ratings and open-ended comments focusing on general and personal interests ("Tutoring is good for my child." "The tutor seemed nice."). These results and the small sample size lead the evaluator to regard the evidence as insufficient or inconclusive at this time.

⁵CREP has developed scannable and online surveys. For more information, go to <http://crep.memphis.edu>. Also, samples of selected surveys may be seen in Appendix A. Sample Instruments/Tools.

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	Usually one coordinator per district; if more, the evaluator may want to include all.	Survey of district coordinators with follow-up phone interview if needed
Principals or school representatives	At least one representative at each SES school, responding for each provider serving students at that school	Survey of principals or school representatives with follow-up phone interview if needed
Teachers	At least 8 but preferably as many as 15–20 teachers per provider, depending on the number of students served	Survey of teachers
Parents	At least 5 but up to 20 per provider, depending on the number of students served	Survey of parents
Providers	One representative from each provider	Survey of providers
Students (when feasible)	From 5–10 students per provider, depending on the number of students served	Survey of students

VI. PROVIDER EVALUATION RUBRIC AND DECISION TREE

A state’s overall evaluation of a 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 or interviews with students and parents of students who receive SES.
3. **Service delivery and compliance.** Measured through surveys or 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 when conducting a summative evaluation of each provider. Based on the summative evaluation results, each provider’s status for the following year can be determined. Figure 1 provides a graphic representation of a summative evaluation.⁶

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 and 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

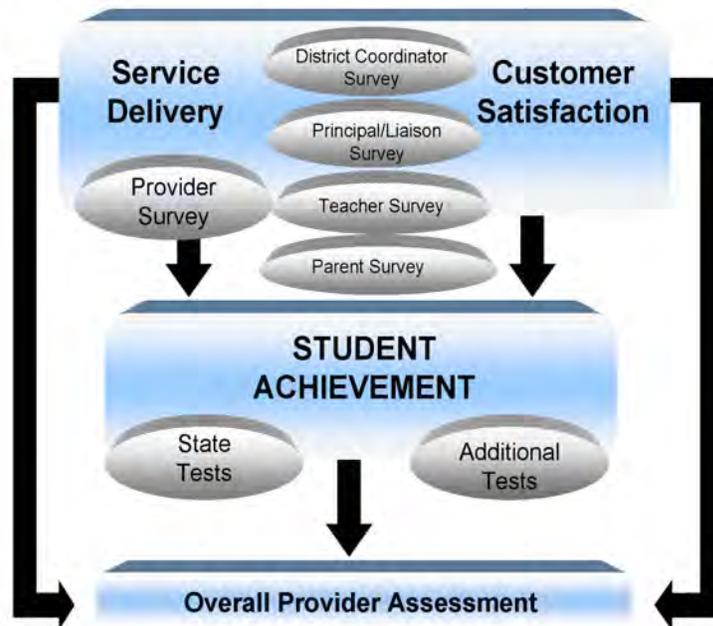
⁶This model is also used in CREP’s comprehensive SES Evaluation/Monitoring Plan (see <http://crep.memphis.edu>).

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.XX⁷ 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 the use of effective communications with these groups.

Figure 1. Summative Evaluation Overview

Figure 1. Components of a Comprehensive SES/Evaluation Modeling Plan



Rubric Application

A sample rubric used in a state evaluation conducted by CREP may be viewed in appendix B.⁸ In addition to evaluating student achievement, this particular state established supplemental core categories, including: Communication, Instructional Plans, Local and State Standards, Special ED/ELL Students, and Overall Performance. For each category, the evaluator synthesizes the available data to select one of five ratings:

- * Insufficient information
- * Below standards
- * Marginal quality
- * Acceptable
- * Above standards

⁷States will need to determine the effect size.

⁸CREP has developed an electronic template and scoring system that automatically records the ratings and stores them in a database for further analyses and possible comparison to state or national norms (depending on SEA preference).

To ensure reliability of findings, each evaluation is reviewed by at least two independent raters. Where there is disagreement, a consensual rating is derived.

Decision Tree Application

Once the rubric ratings have been completed, the culminating step is determining each provider's status for the following year. Appendix C illustrates an example of a generic decision tree that CREP has devised for such purposes, with status category descriptions illustrated in appendix D. The assumption was that individual states could use the generic framework as a starting point but make adaptations to fit local policies or preferences.

The decision tree establishes a hierarchy of status levels, ranging from “full standing” to “removal.” Intermediate levels from highest to lowest include “satisfactory standing,” “probation I,” and “probation II.” The decision logic is designed to give the strongest positive weight to providers that demonstrate “clear” effectiveness in raising achievement. Conversely, the strongest negative weight is assigned to clear failure to raise achievement but also to serious compliance violations.⁹ When student achievement outcomes for a provider are indeterminate (which may be quite common, particularly in the first few years of data analysis, due to small sample size), the decision will assign more weight to implementation quality (e.g., consumer satisfaction variables) and across years, to implementation improvement. This particular model, therefore, adopts a “developmental” orientation that attempts to foster accountability and continuous improvement for providers that cannot yet demonstrate reliable effects on achievement (mostly due to statistical or sampling constraints) but appear to be offering beneficial services.

VII. COMMUNICATING EVALUATION PLAN AND RESULTS

Two important areas of communication for states involve communicating the evaluation plan and expectations to providers and other stakeholders prior to the school year. In fact, states might consider requiring providers to acknowledge formally that they are both aware of evaluation requirements and will cooperate with data collection as needed.

Once the evaluation is completed, 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 will typically give 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.

⁹Criteria for establishing “clear” (statistically significant and demonstrable) criteria for achievement effects and defining the seriousness of compliance violations would generally be decided by individual states. Guidance or technical assistance is available from the authors.

VIII. CONCLUSION

The evaluation of SES providers is a relatively 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 Guidebook 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 tradeoff 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 Guidebook 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., Regression Analysis) is broadly applied across the state, and (b) a highly rigorous design (e.g., Matched Student Pairs Analysis) 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 the effectiveness of evaluation practices and methods that are used to assess the strengths and weaknesses of programs, policies, personnel, products, and organizations.

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.

Center on Innovation & Improvement (CII)

www.centerii.org

CII is one of five content centers in ED's system of comprehensive centers, which also includes 16 regional centers. CII's Web site provides information on SES and other areas of school improvement and innovation.

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. CCSSO's Web site offers guidance and resources on afterschool programs and SES.

National Reporting System (NRS) for Adult Education

www.nrsweb.org

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

Office of Innovation and Improvement

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

This office of ED 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 ED'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.

X. ACKNOWLEDGEMENTS AND DISCLAIMERS

This document was collaboratively developed by the Center for Research in Educational Policy at the University of Memphis and the American Institutes for Research and updated with support from the Center on Innovation and Improvement. We would like to express our appreciation to those who helped contribute to the original edition of this Guidebook, 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, formerly of the Education Leaders Council. We also extend thanks to the SESQ Center Coordinating Group for their valuable input. The development of the original Guidebook was supported through a grant from the office of Innovation and Improvement of 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 AIR

The American Institutes for Research® (AIR®) is one of America's leading educational research and evaluation organizations. Since its founding in 1946 as an independent, not-for-profit corporation, AIR has provided school districts, states, and the federal government with consulting services in the areas of research, analysis, operations, technical assistance, assessment, and strategic communications. AIR is headquartered in Washington, DC, with 10 regional offices throughout the United States and 12 international offices. AIR currently has over 1,100 employees, more than half of whom work in the area of education, and many of whom are former school teachers, principals, and central office and state department of education employees.

AIR staff have extensive expertise in evaluating the implementation and effectiveness of supplemental educational services (SES) and providing technical assistance to those who are responsible for SES, as well as expertise in engaging and empowering parents. From 2003–2005, AIR's Supplemental Educational Services Quality Center (SESQ Center—www.tutorsforkids.org), funded by the Office of Innovation and Improvement of the U.S. Department of Education (ED), supported effective implementation of SES. Through the SESQ Center, AIR provided research, technical assistance, products, presentations and consulting services to (a) engage families to build informed demand for SES, (b) increase the supply of high-quality SES providers, and (c) provide information and improved coordination among national organizations that support the implementation of SES.

Currently, AIR is assisting states with SES evaluation and outreach efforts through its work on several Regional Comprehensive Assistance Centers, funded by ED. In addition, AIR has partnered with the Center for Research in Educational Policy to directly help several states evaluate SES providers. AIR is also working on the National Longitudinal Study of No Child Left Behind (NLS-NCLB), and the Study of State Implementation of Accountability and Teacher Quality Under NCLB (SSI-NCLB). These linked studies are both 4-year, longitudinal evaluations. NLS-NCLB focuses on the implementation of NCLB at the district and school level, while SSI-NCLB collects interview and extant data from states. Both studies have collected the most complete, nationally representative data that currently exists on the implementation of SES.

Web site: www.air.org

Phone: 202-403-5000



About the Center for Research in Educational Policy

The Center for Research in Educational Policy (CREP) is a University of Memphis-based research center and State of Tennessee Center of Excellence whose mission is to implement a research agenda associated with educational policies and practices in pre-K–12 schools and to disseminate research findings so that they inform decisions made by educational practitioners and policy makers. CREP has gained national recognition for its contribution to discussions of issues such as reform of teacher education, educational equity, educational technology, school reform and restructuring, urban and multicultural education, interventions for at-risk students, and using formative evaluation methods for school improvement decision making. In particular, CREP is one of the most active and expert organizations in guiding national policies and state practices for conducting SES evaluations.

One major focus for CREP is to plan and conduct high-quality studies individually and in collaboration with peer organizations to determine “what works” in schools across the nation. In this regard, CREP has served as a consultant or direct contributor to research standards designed to address the goals of No Child Left Behind (NCLB), as proposed by the National Academy of Sciences and the What Works Clearinghouse.

In addition to this research focus, another priority for CREP is assisting schools in building their capacity and ability to make data-driven decisions for the purpose of attaining NCLB Adequate Yearly Progress standards. This emphasis was reflected during the 2005–2006 school year in which approximately 2,000 schools across the nation used tools and strategies CREP developed as part of its Formative Evaluation Process for School Improvement (FEPSI). These tools address critical school improvement factors such as teaching methods and quality, school climate, teacher support, technology usage, program implementation and student achievement. Exemplary projects in which CREP has been recently involved include:

- * Evaluator of Reading First in Tennessee and Louisiana
- * Evaluator of charter schools in Tennessee
- * Evaluator of educational technology initiatives in Tennessee, Michigan, and Kentucky
- * Evaluator of the New Leaders for New Schools Principal Development Program
- * Evaluator or Co-Evaluator of Comprehensive School Reform in Georgia, Tennessee, West Virginia, and Kentucky
- * Evaluator of SES in Tennessee, Louisiana, Virginia, Kentucky, Illinois, and Texas
- * Evaluator of Transition to Teaching Program in Tennessee

Web site: www.crep.memphis.edu

Phone: 866–670–6147

CENTER ON INNOVATION & IMPROVEMENT

Twin paths to better schools

About the Center on Innovation & Improvement

The Center on Innovation & Improvement (CII) is one of five content centers in the U. S. Department of Education's system of comprehensive centers, which also includes 16 regional centers. CII's goal is to provide information, tools, and training to the regional comprehensive centers so they can help states improve student learning by:

- * Assisting districts and schools in their improvement
- * Implementing and supporting the choice provisions of No Child Left Behind
- * Increasing the number and quality of charter schools
- * Ensuring equitable participation of private school students and teachers in federal programs for which they are eligible.

The Academic Development Institute (Lincoln, Illinois) and its partners—Temple University Center for Research in Human Development and Education (Philadelphia, Pennsylvania) and Little Planet Learning (Nashville, Tennessee)—operate the CII with an experienced staff, a Scientific Council with distinguished scholars, a Leadership Council of experienced school administrators, and technical advisors with specific expertise.

To support the regional comprehensive centers' work with states to improve and expand supplemental educational services (SES), CII: (a) collaborates with the Center for Research in Educational Policy and the American Institutes for Research to provide up-to-date information on evaluation of SES services and systems, (b) assists regional centers in providing technical assistance on SES to state education departments, (c) provides training and technical assistance in a program model to engage community-based organizations and faith-based organizations in outreach to parents about SES, and (d) provides information about SES on its Web site, including research, reports, tools, state policies, state programs, and state progress.

CII's other areas of priority are restructuring and state systems of support for schools in improvement. Information about these priority areas and other CII projects may be found at CII's Web site.

Web site: www.centerii.org

Phone: 217-732-6462

Supplemental Educational Services Principal/Site Coordinator Questionnaire

2005[©] Potter, A., McDonald, A.J., & Ross, S.M.
Center for Research in Educational Policy, The University of Memphis.
All Rights Reserved.

DIRECTIONS

USE NO. 2 PENCIL ONLY

MAKE DARK MARKS

EX ○ ○ ○ ○ ○

ERASE COMPLETELY TO CHANGE

District Name:	<input type="radio"/> Davidson County	<input type="radio"/> Hamilton County	<input type="radio"/> Knox County
	<input type="radio"/> Fayette County	<input type="radio"/> Hardeman County	<input type="radio"/> Memphis City
School Name:	<input style="width: 100%;" type="text"/>		
Name and Title of Person Completing this Survey:	<input style="width: 100%;" type="text"/>		

<p>Provider Name: <input style="width: 90%;" type="text"/></p> <p>Provider Code:</p> <table style="width: 100%; text-align: center;"> <tr> <td style="width: 10px;">1</td><td style="width: 10px;">2</td> </tr> <tr> <td><input type="radio"/></td><td><input type="radio"/></td> </tr> </table> <p style="font-size: small; background-color: black; color: white; padding: 2px;">Instructions: Please indicate the particular provider to which this survey pertains. Use only one 2-digit provider code from the 50 listed on the back of this page. A separate survey is needed for each provider.</p>	1	2	<input type="radio"/>	<p>How many of your students did this provider serve this school year?</p> <p><input type="radio"/> 1-10</p> <p><input type="radio"/> 11-20</p> <p><input type="radio"/> 21-30</p> <p><input type="radio"/> Over 30</p>	<p>What was the start date of provider services?</p> <table style="width: 100%;"> <tr> <td style="width: 50%;">Month</td> <td style="width: 50%;">Date</td> </tr> <tr> <td><input type="radio"/> October</td> <td><input type="radio"/> 01</td> </tr> <tr> <td><input type="radio"/> November</td> <td><input type="radio"/> 02</td> </tr> <tr> <td><input type="radio"/> December</td> <td><input type="radio"/> 03</td> </tr> <tr> <td><input type="radio"/> January</td> <td><input type="radio"/> 04</td> </tr> <tr> <td><input type="radio"/> February</td> <td><input type="radio"/> 05</td> </tr> <tr> <td><input type="radio"/> Other _____</td> <td><input type="radio"/> 06</td> </tr> <tr> <td></td> <td><input type="radio"/> 07</td> </tr> <tr> <td></td> <td><input type="radio"/> 08</td> </tr> <tr> <td></td> <td><input type="radio"/> 09</td> </tr> <tr> <td></td> <td><input type="radio"/> 10</td> </tr> </table>	Month	Date	<input type="radio"/> October	<input type="radio"/> 01	<input type="radio"/> November	<input type="radio"/> 02	<input type="radio"/> December	<input type="radio"/> 03	<input type="radio"/> January	<input type="radio"/> 04	<input type="radio"/> February	<input type="radio"/> 05	<input type="radio"/> Other _____	<input type="radio"/> 06		<input type="radio"/> 07		<input type="radio"/> 08		<input type="radio"/> 09		<input type="radio"/> 10	<p>In which subjects did your students receive services from this provider?</p> <p><input type="radio"/> Reading/Language Arts only</p> <p><input type="radio"/> Mathematics only</p> <p><input type="radio"/> Both Reading/LA and Math</p> <p><input type="radio"/> Uncertain/Do Not Know</p> <p><input type="radio"/> Other _____</p>																			
1	2																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
<input type="radio"/>	<input type="radio"/>																																														
Month	Date																																														
<input type="radio"/> October	<input type="radio"/> 01																																														
<input type="radio"/> November	<input type="radio"/> 02																																														
<input type="radio"/> December	<input type="radio"/> 03																																														
<input type="radio"/> January	<input type="radio"/> 04																																														
<input type="radio"/> February	<input type="radio"/> 05																																														
<input type="radio"/> Other _____	<input type="radio"/> 06																																														
	<input type="radio"/> 07																																														
	<input type="radio"/> 08																																														
	<input type="radio"/> 09																																														
	<input type="radio"/> 10																																														

Are you employed by the provider for which you are completing this survey? Yes No

Indicate your response to each of the following items.

		Frequently	Occasionally	Not at all	Don't know
How often does the provider...					
Communicate with you during the school year?	<input type="radio"/>				
Meet the obligations for conducting tutoring sessions?	<input type="radio"/>				
Communicate with teachers during the year?	<input type="radio"/>				
Communicate with parents during the year?	<input type="radio"/>				

		Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
The provider...						
Adapted the tutoring services to this school's curriculum.	<input type="radio"/>					
Integrated the tutoring services with classroom learning activities.	<input type="radio"/>					
Aligned their services with state and local standards.	<input type="radio"/>					
Offered services to Special Education and ESL students.	<input type="radio"/>					
Complied with applicable federal, state, and local laws.	<input type="radio"/>					

Overall assessment:

I believe the services offered by this provider positively impacted student achievement.	<input type="radio"/>				
Overall, I am satisfied with the services of this provider.	<input type="radio"/>				

Additional comments can be provided on the back on this form.

State of Tennessee Supplemental Educational Services Parent Questionnaire

2005 © A. Potter, McDonald, A.J., & Ross, S.M.
Center for Research in Educational Policy, The University of Memphis.
All Rights Reserved.

Instructions:



According to school records, your child is receiving Supplemental Educational Services. Please fill in the name of your child's provider, district, and school. A list of provider names is on the back of this page.

Provider Name:

District Name:

School Name:

Please answer the following survey questions about the services your child has received.

What was the date the tutoring services started this year?

- October 2005
- November 2005
- December 2005
- January 2006
- February 2006
- Other _____ 2006

In which subjects did your child receive services?

- Reading/Language Arts only
- Mathematics only
- Both Reading/Language Arts and Math
- Uncertain/Do Not Know
- Other _____

Indicate your response to each of the following items.

Frequently
 Occasionally
 Not at all
 Don't know

How often does the provider...

Talk to me about my child's progress?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Talk to my child's teachers about his/her progress?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Send letters or notes home to me about my child's progress?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Help my child with subjects s/he is working on in the regular school classroom?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Start and end the tutoring sessions at the scheduled time?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Indicate how much you agree or disagree with each of the following items about the provider.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

I believe that the services offered have helped my child's achievement.	<input type="radio"/>				
Overall, I am pleased with the services that my child received.	<input type="radio"/>				

Indicate how much you agree or disagree with each of the following items about the School District.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

I was given information about my child's rights under the No Child Left Behind law.	<input type="radio"/>				
I was given enough time to decide which service provider I wanted for my child.	<input type="radio"/>				
I am pleased with the way my school district helped me obtain Supplemental Educational Services for my child.	<input type="radio"/>				

Additional comments can be provided on the back on this form.

Do not fold this questionnaire!

Do not fold this questionnaire!

Tennessee Supplemental Educational Services Provider Questionnaire



Please make copies of this form and complete a separate survey for each district with which you have/had a contract to serve students with Supplemental Educational Services in 2005-2006.

I. Supplemental Educational Services Provider Information

Provider Name: <input style="width: 95%;" type="text"/>	Please describe the format of your services by completing the following items:
Contact Person: <input style="width: 95%;" type="text"/>	Program duration (e.g., 10-weeks; academic year) <input style="width: 95%; height: 25px;" type="text"/>
Address: <input style="width: 95%;" type="text"/>	Weekly duration (e.g., each student attends twice per week for two hours total) <input style="width: 95%; height: 25px;" type="text"/>
Telephone: <input style="width: 95%;" type="text"/>	Setting (e.g., school, your building) <input style="width: 95%; height: 25px;" type="text"/>
E-Mail: <input style="width: 95%;" type="text"/>	Format (e.g., small group, individual) <input style="width: 95%; height: 25px;" type="text"/>
District Served: <input style="width: 95%;" type="text"/>	If groups are used, what is the average size? <input style="width: 95%; height: 25px;" type="text"/>
Name of School served in this district: <input style="width: 95%;" type="text"/>	Is transportation provided to students? If so, by whom (you, district, school, etc.)? <input style="width: 95%; height: 25px;" type="text"/>

What is your general instructional approach during the tutoring sessions: (e.g., students work one-on-one with the tutor and then complete practice/review exercises on a computer)?

Is there anything else that you would like to say regarding the format/process of your services?

Qualifications of tutors (certified teachers, teaching aide, training, etc):

Are background checks done on tutors or other personnel that come into contact with students? (If yes, please describe).

Form completed by (if other than contact person):
Address, Telephone, and Email

**Tennessee Supplemental Educational Services Provider Questionnaire
(continued)**

II. Information Regarding Students You Served in this District

1. Complete the table below based on data for this district. If unable to respond to a particular section, provide an explanation in the space provided below the table.

Grade Level	# of Total Students Served	# of Special Ed Students Served	# of Limited English Proficiency (LEP) Students Served	# of Students Who Achieved Their Specific Achievement Goals	# of Students Who Made Progress In Their Specific Achievement Goals	# of Students Who Showed No Improvement In Their Specific Achievement Goals	% of Tutoring Sessions Attended by Students

Notes: (Use the space below to comment on missing data or any special circumstances relevant to the above outcomes.)

2. For those students who did not achieve their stated goals, explain why this did not happen, either generally or on an individual basis.

**Tennessee Supplemental Educational Services Provider Questionnaire
(continued)**

3. Indicate the number of students served in each curriculum area in which your organization offered services in this district. Indicate numbers in all that apply.

<input type="checkbox"/> Reading	<input type="checkbox"/> Writing	<input type="checkbox"/> Math	<input type="checkbox"/> Other _____
K _____	K _____	K _____	K _____
1 _____	1 _____	1 _____	1 _____
2 _____	2 _____	2 _____	2 _____
3 _____	3 _____	3 _____	3 _____
4 _____	4 _____	4 _____	4 _____
5 _____	5 _____	5 _____	5 _____
6 _____	6 _____	6 _____	6 _____
7 _____	7 _____	7 _____	7 _____
8 _____	8 _____	8 _____	8 _____
9 _____	9 _____	9 _____	9 _____
10 _____	10 _____	10 _____	10 _____
11 _____	11 _____	11 _____	11 _____
12 _____	12 _____	12 _____	12 _____

III. Provider Perceptions and Activities

(Indicate your response to each of the following items as they apply to your staff.)

Don't know
 Not at all
 Occasionally
 Frequently

1. Tutors communicated with teachers regarding progress of their student(s).

If applicable, briefly describe the process for such communications:

2. Tutors communicated with parents/guardians regarding their child's progress.

If applicable, briefly describe the process for such communications:

3. Tutors adapted the supplemental services to each school's curriculum

If applicable, briefly describe the process for such adaptations:

Tennessee Supplemental Educational Services Provider Questionnaire (continued)

III. Provider Perceptions and Activities (Continued)

Don't Know
Not at all
Occasionally
Frequently

4. Tutors aligned the supplemental services with the state academic content and achievement standards.

If applicable, briefly describe the process for such alignments:

5. Tutors integrated the tutoring services with classroom learning activities.

If applicable, briefly describe the process for such processes:

6. Tutors showed their lesson plans or materials used for tutoring to the homeroom/subject teacher of each child they worked with.

If applicable, briefly describe the process for such communications:

7. Tutors gave instruction to students with disabilities, consistent with their Individualized Education Plans or Individualized Services under Section 504 of the Rehabilitation Act.

If applicable, briefly describe the process for such activities:

8. Tutors protected from public disclosure the identities of all students served and all students eligible for services.

If applicable, briefly describe the process for such activities:

**Tennessee Supplemental Educational Services Provider Questionnaire
(continued)**

III. Provider Perceptions and Activities (Continued)

Don't Know
Not at all
Occasionally
Frequently

9. Tutors give appropriate instruction to English Language Learners if it is needed.

If applicable, briefly describe the process for such activities:

10. Tutors use appropriate timetables for improving each student's achievement.

If applicable, briefly describe the process for such activities:

Based on your perceptions and experiences, rate the degree of satisfaction with each of the following areas. Use the accompanying "Comments" section to elaborate if desired, especially where "highly dissatisfied" or "highly satisfied" is indicated.

Highly Dissatisfied
Dissatisfied
Don't Know
Satisfied
Highly Satisfied.

11. Parent cooperation/involvement

Comments:

12. Student attendance

Comments:

Tennessee Supplemental Educational Services Provider Questionnaire
(continued)

Highly Dissatisfied
Dissatisfied
Don't Know
Satisfied
Highly Satisfied

13. Student attitudes (e.g., cooperation, motivation)

Comments:

14. The ease of developing lessons aligned with the district or school curriculum.

Comments:

15. Teacher cooperation/involvement

Comments:

16. District cooperation/involvement

Comments:

17. Success at raising student achievement to desired levels

Comments:

**Tennessee Supplemental Educational Services Provider Questionnaire
(continued)**

Overall Impressions

19. What was the most positive outcome or aspect of your work with the district this year?

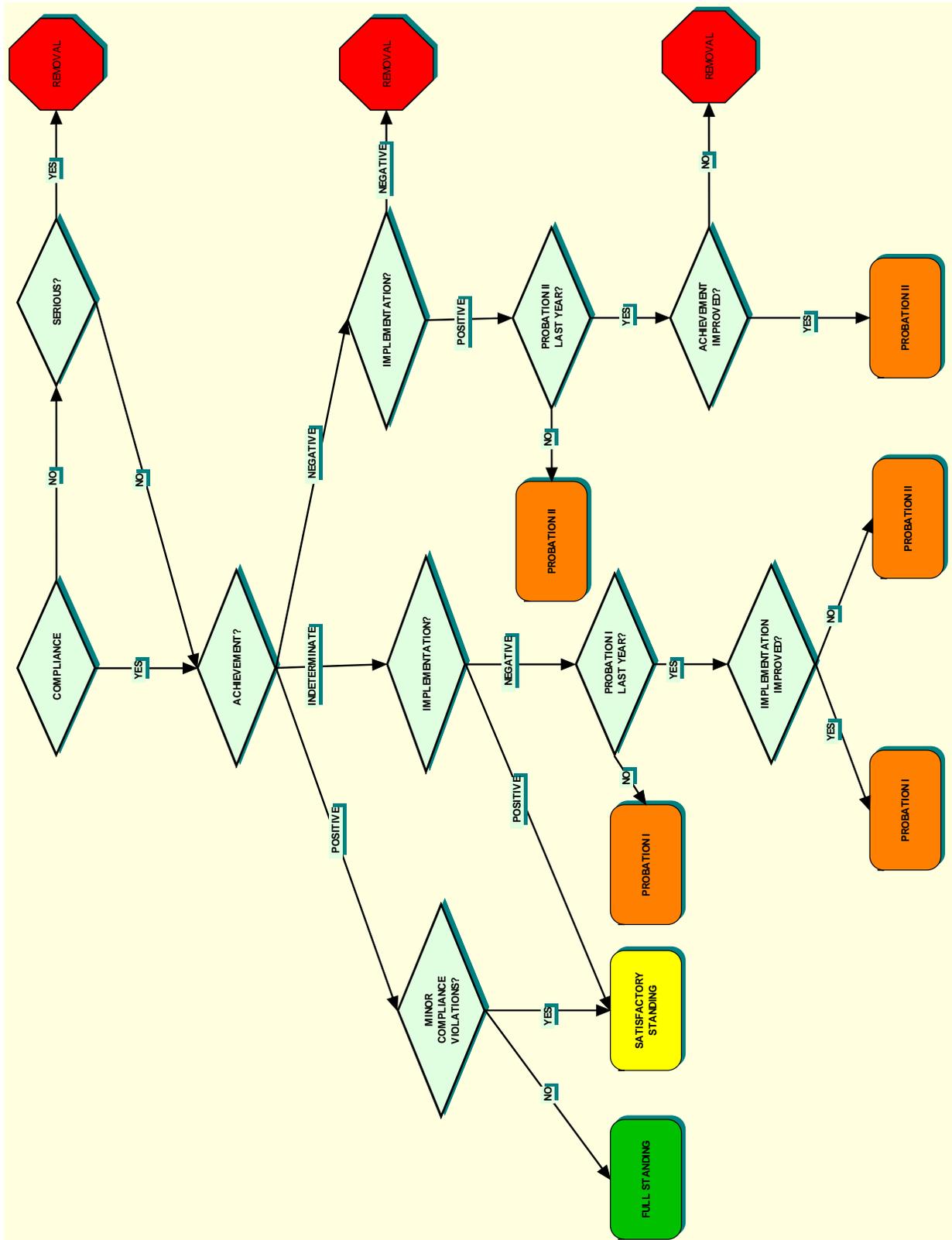
20. What was the most negative aspect or area in need of improvement regarding your work with this district this year?

21. Additional Comments/Recommendations

Appendix B. Rubric of Overall Provider Effectiveness

Outcome	Insufficient Information	Below Standards	Marginal Quality	Acceptable	Above Standards
1. Student Achievement	There is insufficient information available to determine student achievement outcomes.	Students have not shown gains related to tutoring received from service providers.	About half of the students have made some gain related to tutoring received from service providers.	There has been some gain for the majority (over 60%) of students related to tutoring received from service providers	The effect size for students in the provider's program is in the top one-third of all the effect sizes demonstrated by providers meeting standards for student achievement.
2. Communication	There is insufficient information available to determine communication outcomes.	Provider has not communicated with the principals, teachers, and parents of students served.	There has been limited communication throughout the year between the provider and at least two of the following: principals, teachers, and parents.	There has been some regular communication throughout the year between the provider and the principals, teachers, and parents of students served.	There is an ongoing and sustained system of communication between the provider and the school-level educators as well as parents of students served.
3. Instructional Plans	There is insufficient information available to determine instructional plans of the provider.	Provider does not plan instruction explicitly geared to student needs or to reinforce their regular academic program.	Provider is in the planning stages of gearing instruction to student needs, and reinforcing the regular academic program.	Provider has made some attempt with the majority of students to plan instruction explicitly geared to student needs and to reinforce the regular academic program	Provider instructional plans are explicitly geared to the needs of most or all students and reinforce the regular academic program.
4. Local and State Standards	There is insufficient information available to determine alignment with the local and state standards.	None of the instructional plans used by the provider are aligned with local and state academic standards for students.	Provider is in the process of aligning instructional plans with local and state academic standards for students.	Some of the instructional plans used by the provider are presently aligned with local and state academic standards for students.	Most or all of the instructional plans are presently aligned with local and state academic standards for students.
5. Special Ed/ELL Students	There is insufficient information available to determine special ed/ELL student outcomes.	Provider does not offer accommodations for addressing the needs of special ed or ELL students.	Provider has made limited accommodations for addressing the needs of special ed and ELL students.	Provider has made some accommodations for addressing the needs of special ed and ELL students.	Provider offers appropriate services, if needed, to special education and ELL students.
6. Provider Overall	There is insufficient information available to determine provider overall outcomes	There is overall dissatisfaction with the provider at the district and school levels.	There is more dissatisfaction than satisfaction with the provider at the district and school levels.	There are mixed but mostly positive reactions about the provider at the school and district levels.	There is overall satisfaction with the provider at the district and school levels.

Appendix C. Decision Tree for SES Providers



Appendix D. Descriptions of Provider Status Categories

Full Standing:	The provider has demonstrated positive achievement effects, and is fully approved to continue services in the following year.
Satisfactory Standing:	<p><i>Remedy requirement: None</i></p> <p>The provider has demonstrated “indeterminate” achievement effects due to insufficient data or equivocal effect sizes, but has shown positive implementation outcomes (customer satisfaction, service delivery, compliance).</p>
OR	<p>The provider has demonstrated positive achievement effects but has minor compliance violations.</p> <p><i>Remedy Requirement: Submit improvement plan for correcting any compliance violations OR for improving student achievement where lack of positive evidence is due to equivocal outcomes (neither sufficiently positive nor negative effect sizes) not to insufficient sample size.</i></p>
Probation I:	<p>The provider has demonstrated “indeterminate” achievement effects due to insufficient data or equivocal effect sizes, and has weak or negative implementation outcomes.</p> <p><i>Remedy Requirement: Submit improvement plan for improving implementation quality, any compliance violations, and student achievement where lack of positive evidence is due to equivocal outcomes (neither sufficiently positive nor negative effect sizes) not to insufficient sample size.</i></p>
Probation II:	<p>The provider has demonstrated “negative” achievement effects but has positive implementation outcomes.</p> <p><i>Remedy Requirement: Submit improvement plan for improving student achievement.</i></p> <p>OR,</p> <p>The provider was in Probation I status last year, but failed to improve implementation.</p> <p><i>Remedy Requirement: Submit improvement plan for improving implementation and student achievement.</i></p>
Removal:	<p>The provider has demonstrated “negative” achievement effects and has weak or negative implementation outcomes.</p> <p>OR,</p> <p>The provider was in Probation II status last year, and although still demonstrating positive implementation outcomes, has failed to improve student achievement.</p> <p><i>Remedy Requirement: None, but should the current year achievement outcomes (to become available in the fall of the next school year) show positive effects, the provider may apply for reinstatement after being removed for one year.</i></p>

NOTE: Separate status determinations will be made for each subject tutored. Thus, it is possible that a provider offering tutoring in both reading and mathematics could be classified as Full Standing in one subject but Removal in the other.

For more information please contact:

CENTER ON INNOVATION & IMPROVEMENT
121 N. Kickapoo Street
Lincoln, Illinois 62656
217-732-6462
217-732-3696
www.centerii.org

