

---

# **Scientifically based research and the Comprehensive School Reform (CSR) Program**

---

**August 2002**

Comprehensive School Reform Program Office

Office of Elementary and Secondary Education  
U.S. Department of Education

This document may be duplicated for dissemination.

---

---

## **Introduction**

The primary responsibility of schools undertaking comprehensive school reform is creating programs that result in improved student achievement. One of the most important tasks in this process is choosing highly effective reform strategies, methods, and programs, those that are grounded in scientifically based research. This guide is designed to help school staffs and those with whom they work to increase their understanding of what scientifically based research is, and to use that understanding to assess the quality, relevance and usefulness of the research they examine.

## **How has the reauthorization of the ESEA raised the standard for evidence of effectiveness?**

Determining that reform strategies are effective before implementing them is not only common sense, but is also one of the requirements of the Comprehensive School Reform program. The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA) requires that only those strategies and methods proven effective by the standard of scientifically based research should be included in school reform programs. As defined in the ESEA, “scientifically based research” emphasizes experimental and quasi-experimental studies that are systematic, empirical, well-designed, replicable, and have been accepted by independent reviewers. (See Appendix 2 for complete definition.)

Prior to the 2001 reauthorization, CSR-supported programs were required to use “*innovative* strategies and proven methods for student learning, teaching, and school management . . . based on *reliable research* and effective practices, and . . . replicated successfully in schools with diverse characteristics.” The new law requires comprehensive reform programs that “employ *proven* strategies and *proven* methods for student learning, teaching, and school management that are based on *scientifically based research* and effective practices and have been replicated successfully in schools.”

In addition, the law now addresses the school’s comprehensive program as a whole. It requires of schools a “program that—has been found, through scientifically based research, to significantly improve the academic achievement of students participating in such program as compared to students in schools who have not participated in such program *or* has been found to have strong evidence that such program will significantly improve the academic achievement of participating children.”

---

Clearly, the new law sets a higher bar for evaluating the research base of both the comprehensive program and its individual strategies and methods. Both must reflect “scientifically based research” when such research exists. However, in some cases, research that meets these criteria is minimal or non-existent. For instance, little or no scientifically based research has been conducted that examines the programmatic impact of the interaction of the eleven required components of the CSR program. In this case, school leaders will need to rely on the best available empirical evidence and some degree of professional judgment in creating their programs. As the quantity, quality and availability of empirical, randomized studies increases, schools will be able to make a stronger connection between their design decisions and the evidence of “what works.”

### **CSR components**

To receive CSR funding, schools must implement a comprehensive school reform program that:

1. **Uses proven strategies and methods for learning, teaching, and school management based on scientifically based research and effective practices, and used successfully in multiple schools**
2. Integrates a comprehensive design with aligned components focused on helping students meet standards and addressing needs identified in a school needs assessment
3. Provides high quality, ongoing professional development
4. Includes measurable goals and benchmarks for student academic achievement
5. Has the support of staff within the school
6. Provides support for all faculty and staff
7. Provides for parental and community support and involvement
8. Uses high quality, external technical support and assistance from an experienced provider
9. Includes a plan for the annual evaluation of the implementation of the reform program and the outcomes achieved
10. Identifies other resources to support the reform effort
11. **Has been found through scientifically based research to significantly improve student academic achievement, or has shown strong evidence that it will.**

Components one and eleven identify a legislative (and common sense) standard of scientifically based research. The first component requires that schools apply that standard to the instructional<sup>1</sup> strategies contained within their comprehensive school reform programs. Component eleven, on the other hand, requires either scientifically based research or “strong evidence” of effectiveness; this standard applies to the remaining components, those that make up the rest of the school’s comprehensive program. For these components, little or no scientifically based research exists, so practitioners must rely on a less

---

<sup>1</sup>Management strategies also are named as part of Component 1 but, given the very limited nature of the research of the impact of management strategies on achievement, we recommend focusing on instructional practices here.

---

rigorous standard. Findings from research studies that fulfill most, but not all of the requirements of scientifically based research as defined in the law demonstrate “strong evidence.”

### **Standards for identifying effective approaches**

The reauthorized ESEA stresses that schools should review scientifically based research to determine that the reform approaches they are considering are likely to have **a positive impact on student achievement**. The questions on the following pages and the accompanying *Guidelines for Judging the Quality of a Study* can guide that review process. Although information about “what works” ranges from folklore to case studies to studies with randomized trials, this publication focuses on identifying the “silver” and “gold” research standards, the highest levels of scientifically based research as it is defined in the law.

### **Research consumer questions:**

Finding evidence of effects on student achievement is important, but in order to gain a broad understanding of the potential usefulness of a reform, consumers are advised to act as discriminating “research consumers,” examining the research from three perspectives:

1. the **theoretical base** of the reform practice or program
2. **implementation and replicability** information
3. **evidence of effects on student achievement**

Some questions designed to help school staffs think about these issues are provided on the next page. Additional resources are identified in Appendix 1.

---

**Question 1:****Is there a theoretical base for the practice or program being considered?**

<b>Questions about the theoretical base</b>	<b>Judging quality of the theoretical base</b>
<ul style="list-style-type: none"><li>▪ What are the ideas behind this practice or program?</li><li>▪ What are its guiding principles?</li> <li>▪ How does it work?</li> <li>▪ Why does it work?</li></ul>	<ol style="list-style-type: none"><li>1) Is there a clear, non-technical description of the central idea and goals of the practice or program?</li> <li>2) Is there a clear description of the instructional activities that are central to this program or practice?</li> <li>3) Is the practice clearly tied to an established learning theory, e.g. child development or language acquisition?</li></ol>

**Question 2:****Is there evidence that this practice or program has been successfully implemented and has produced positive outcomes in a variety of situations? Has it been successful in a context similar to that of the school considering this practice?**

<b>Questions about implementation and replicability</b>	<b>Judging quality of implementation and replicability</b>
<ul style="list-style-type: none"><li>▪ Has this program or practice been widely used?</li> <li>▪ Where is this reform likely to work?</li> <li>▪ Under what circumstances is it most effective?</li></ul>	<ol style="list-style-type: none"><li>1) How many schools have used this practice or program?</li> <li>2) Did the schools using it fully implement the practice or program?</li> <li>3) In what settings has it been implemented?</li> <li>4) Has improved student achievement been convincingly demonstrated in a variety of settings?</li></ol>

**Question 3: Is there evidence that this practice or program has a significant positive effect on student achievement?**

<b>Judging the quality of the evidence</b>		
<b>Question about individual practices or programs</b>	<b>Scientifically based research</b>	<b>Developing toward scientifically based research</b>
<p>Is there evidence based on rigorous research showing that this practice and/or program improves student achievement?</p>	<p>For each practice or program identified:</p> <ol style="list-style-type: none"> <li>1) Are there studies looking at the impact on students of that practice or program?</li> <li>2) Are those studies of high quality? (See Guidelines for Judging Quality of Study, pp. 7-10).</li> <li>3) Are there at least 5 high quality studies?</li> <li>4) Do 4 of the 5 high quality studies show that the practice improves student achievement?</li> <li>5) If yes, are the findings significant in 3 of those 4 studies? (See p. 11).</li> </ol>	<p>For each practice or program identified:</p> <ol style="list-style-type: none"> <li>1) Are there studies looking at the impact on students of that practice or program?</li> <li>2) Are those studies of reasonable quality? (See Guidelines for Judging Quality of Study, pp. 7-10).</li> <li>3) Are there at least 5 high or reasonable quality studies?</li> <li>4) Do 4 of the 5 reasonable quality studies show that the practice improves student achievement?</li> <li>5) If yes, are the findings significant in 3 of those 4 studies? (See p. 11).</li> </ol>
	<p>If the answer to all of these questions is YES, there is scientifically based research regarding this practice or program.</p> <p style="text-align: center;"><b>The “Gold” Standard</b></p>	<p>If the answer to all of these questions is YES, there is “strong evidence” regarding this practice or program, even though the research on which it is based did not meet all the requirements of scientifically based research as defined in the law.</p> <p style="text-align: center;"><b>The “Silver” Standard</b></p>

---

## **Guidelines for judging the quality of a study**

The criteria for judging the quality of research studies are contained in the definition of scientifically based research in section 9101(37) of the reauthorized Elementary and Secondary Education Act (ESEA). Although there is no universally accepted standard, for the purposes of this publication, a **high quality study** meets *all* of the criteria described below. A **reasonable quality study** meets *all but one* of the criteria. For example, a reasonable quality study might be *systematic, empirical,* and use *rigorous data analysis on reliable and valid data*, but might use a longitudinal study *design* that does not involve random assignment to study groups or statistical controls on background characteristics.

### **Criteria 1: Systematic and empirical**

High quality research is carried out in a manner that is consistent, disciplined, and methodical—not sloppy or haphazard. Such research shows evidence of careful planning and keen attention to detail. Empirical research is grounded in data drawn from observation or experiment; the claims being made are supported by measurable evidence, not opinion or speculation.

When evaluating research, consider the following:

- Does the research have a sound theoretical foundation?  
(See Research Consumer Questions, pp. 4 and 5.)
- Were the data obtained using observation or experimentation?
- Were the data collected from *all* appropriate groups of respondents and not just from certain groups? For example, does a school reform program that claims to benefit all students include special education students in its research? If the research uses test results for a given school, did all of the students in the school take the test?
- Were the data observed or collected from multiple subjects (teachers, students, schools, etc.)?
- Are the research findings supported by measurable evidence?

---

## **Criteria 2: Rigorous data analysis**

Even the highest quality data are of little value unless analyzed thoughtfully and carefully. The definition of scientifically based research requires that data collected must be analyzed using methods that are appropriate for the task, and adequate to test the stated hypotheses and justify the general conclusions drawn. Failure to apply appropriate methods could produce inaccurate or misleading findings.

Some key questions to consider about the data analysis include the following:

- Does the research test the stated hypothesis, and do the findings justify the general conclusions drawn?
- Does the research report the sample size and the statistical procedures used?
- Do the researchers analyze the data in a manner appropriate to the research question of interest? Are the statistical procedures used adequate for answering the research question?
- Do the analysis methods correspond to the structure of the data? Does the analysis account for the complexities of the data? for missing data? for unique groupings? for changes in the data over time?  
For example, in school research studies that unfold over time, subjects may drop out of the study (for example, by moving out of a study school). Adequate data analyses address these issues.

## **Criteria 3: Reliable and valid data collection**

High quality data produce accurate and credible findings. Scientifically based research relies on measurements or observational methods that provide reliable and valid data across evaluators and observers and across multiple measurements and observations. Reliability implies that repeated measurements on subjects taken under similar circumstances or over time will produce similar results. If unreliable, the data may hinder the researcher's ability to discern real differences among subjects or programs. To be considered valid, the data collected must measure the outcomes they were designed to measure, (e.g. that student math knowledge is what is being measured, not students' ability to guess test answers). There must be a match between the research question and the observed behavior on which the research findings are based.

Questions about the quality of data collection include the following:

- Was data collection conducted professionally and consistently? For example, was there some system to ensure that different data collectors had the same focus and attention to detail ?  
(E.g. training before data collection or interrater reliability tests)

- 
- Were research biases minimized? Developers of reform models supply a natural example: was the evaluation of the reform model conducted by the model developers or by a third-party, independent evaluator?
  - Does the study look at the appropriate information to address its questions? Are the measures valid? That is, do the measures discussed and analyzed correspond to the concepts being studied?
  - Are the data reliable? Did repeated measurements on subjects taken under similar circumstances produce similar results? Do the data represent counts of actions, records, responses, etc., that directly reflect what the practice or program is supposed to be doing and affecting?

#### **Criteria 4: Strong research design**

Studies must be designed to optimize the investigator's ability to answer the research question or hypothesis.

The following questions are relevant to research design:

- Does the study follow an experimental or quasi-experimental design? That is, are the subjects in the study divided randomly into at least two groups, with at least one group using the practice or program of interest and one group not using it?
- Does the study design contain appropriate controls in order to be able to evaluate the effects of the condition of interest? Were the subjects of the research randomly assigned, or were there other within-condition or across-condition controls as part of the design? (Random assignment of students is a way to ensure that it is the practice or program and not particular student characteristics that are producing the measured results.)
- If subjects are not divided into the groups randomly, are the groups selected to ensure that subjects share similar background characteristics such as economic well-being or previous academic achievement? If not, does the study explain how statistical controls were used to account for these differences in background characteristics of the students in the study? (See criteria 2.)
- Did the research minimize alternative explanations for observed effects?
- Does the study make a determination that the practice or program was used appropriately and fully as intended?

---

### **Criteria 5: Detailed results that allow for replication**

The results of high quality studies are presented in sufficient detail to allow for their replication, or to at least provide opportunities to build systematically on their findings. To increase their usefulness to practitioners, research findings must be reported in a way that makes them easily accessible and understood. The informed lay reader should be able to understand the study's design, methods, and findings.

When evaluating the quality of research reporting, consider the following:

- Are the findings clearly described and reported, free from technical terms and jargon?
- Are the description of the design and the results of the research sufficiently detailed so that replication of the design is possible? For example, do researchers report the sample size (number of people or schools involved) and the statistical procedures used?
- Are the findings presented fairly and objectively?
- Are technical aspects of the study, such as statistical significance or confidence intervals made available and explained? Do the reports supply any supporting technical materials, perhaps in appendices?
- Is the presentation balanced? That is, are shortcomings reported as well as strengths? Were possible explanations provided for findings that ran counter to the researcher's expectations?

### **Criteria 6: Expert Scrutiny**

A strong study should be able to meet criticism by independent, expert reviewers. Peer reviewers, either from scientific journals or from an independent panel of experts in a given field, provide quality control in the form of a rigorous, objective, and scientific review of research. Research consumers can place more confidence in findings that have been subjected to expert review.

When evaluating research, consider the following:

- Has the research been accepted and published by a competitive, peer-reviewed scientific journal, or was it reported only in media such as newspapers, magazines, or trade journals?
- If the work was not published, is there evidence that it was reviewed by independent experts and subjected to external verification? If so, did the reviewers approve the study methodology and interpretation of the findings?

---

### **Significance of effects:**

Before reformers make a final decision about the usefulness of available research findings, they must determine their significance. Even high-quality research studies can produce findings that are not statistically or practically significant. **Significance** is a statistical term that helps readers to understand the likelihood that the findings of a study were the result of the designed intervention and would not be observed independent of that intervention. For practitioners, two standards of significance apply, *statistical* significance and *educational, or practical, significance*.

- For findings to be considered *statistically significant*, researchers should indicate in the study that the findings are “significant” at the .01 level (American Institutes for Research, 1999; Slavin & Fashola, 1998). In the social science research community, a .01 level of significance is highly regarded because it means that the positive results of the study would be the result of some unrelated cause only 1 time in 100 repetitions of the study’s procedures. Thus, a .01 level of significance indicates that the program’s activities have led to the achievement gains reported.
- For the findings to be considered *educationally, or practically, significant*, the effect on student achievement should be large enough to be of practical value. For example, gains on a standardized test should be 10 percentile points to be considered educationally significant (American Institutes for Research, 1999).

A study may have a significant statistical result in that the achievement gains reported in the study are very clearly the result of the reform program. However, the actual gains in achievement may be modest in size; when this is the case, the study may be said to have limited practical significance. In this context, school community members need to decide if the cost of implementing the program is outweighed by the size of the student gains that can be achieved.

### **Putting scientifically based research in the school reform context:**

The comprehensive school reform program usually begins with school staff gathering and analyzing data about their school. Once this process is completed, they identify and prioritize their needs, establish goals, and then begin to research strategies and methods that will help them achieve those goals. The following step-by-step process illustrates the place that scientifically based research occupies in the context of school reform.

<b>Step One:</b>	Schools collect data that gives a complete, accurate picture of the current state of the school.
<b>Step Two:</b>	School community analyzes data to determine and prioritize needs.
<b>Step Three:</b>	After prioritizing needs, schools write goal statements to address prioritized needs.
<b>Step Four:</b>	Schools develop potential strategies to achieve stated goals.
<b>Step Five:</b>	Schools begin topical literature review, identifying research base of potential strategies. (See Appendix 1 for sources of research abstracts.)
<b>Step Six:</b>	Schools use abstracts to determine if the research addresses (1) the theoretical basis of a program or practice, and/or (2) issues of implementation, replicability, and relevance, or (3) evidence of the effects of that program or practice on student achievement. Not all research addresses all three topics, but more than one may be addressed in an individual research study.
<b>Step Seven:</b>	If an examination of the abstract indicates that a study is relevant to the identified needs of the school, the study should be examined in more detail. Schools can use the <i>Standards for Identifying Effective Approaches</i> (pp. 4-6) to make a determination about the practice or program of interest. <i>Guidelines for Judging the Quality of a Study</i> (pp.7-10), and <i>Understanding Significance</i> , (p.11) will provide additional help.
<b>Step Eight:</b>	Schools determine if the examined research is of high quality, replicable, and relevant. The examined research might meet the criteria of scientifically based research, the “strong evidence” criteria, or be at some other point along the spectrum of “evidence based.” In some cases there will be no, or little, high quality evidence that this practice, program or model will increase positive outcomes for students.

The Comprehensive School Reform program is built on the premise that schools can improve when they create a unified, coherent approach to reform, and when they implement practices and programs that are grounded in scientifically based research. Assessing and evaluating the research base of each method and strategy included in the school’s comprehensive design for improvement will improve the quality, implementation, and outcomes of the reforms funded under this program.

---

## References

- American Federation of Teachers. (1997). *Raising student achievement: A resource guide for redesigning low-performing schools*. (AFT Item Number 3780). Washington, DC: AFT.
- American Institutes for Research. (1999). *An Educators' Guide to Schoolwide Reform*. Arlington, VA: Educational Research Service.
- Boruch, R.F., DeMoya, D., and Snyder, B. (in press). The importance of randomized field trials in education and related areas. In F. Mosteller and R. Boruch (Eds.), *Evidence matters: Randomized trials in education research*. Washington, DC: Brookings Institution Press.
- Cook, T.D.. A critical appraisal of the case against using experiments to assess school (or community) effects.. Professor of Sociology, Northwestern University.
- Cook, T.D. (March 1999). Considering the major arguments against random assignment: An analysis of the intellectual culture surrounding evaluation in American schools of education. Northwestern University. Paper presented at the Harvard faculty Seminar on Experiments in Education.
- Cook, T.D., Habib, F., Phillips, M., Settersten, R.A., Shagle, S.C., and Degirmencioglu, S.M. (1999). Comer's School Development Program in Prince George's County, Maryland: A theory-based evaluation. *American Education Research Journal*, 36 (3), pp. 543-597.
- Cook, T.D., Hunt, H.D., and Murphy, R.F. Comer's School Development Program in Chicago: A theory-based evaluation. Institute for Policy Research, Northwestern University.
- Cook, T.D., and Payne, M.R. Objecting to the Objections to Using Random Assignment in Educational Research.
- Cook, T.D. (Fall 2001). Why education researchers reject randomized experiments. Sciencephobia, [Online], Available: <http://www.educationnext.org/>.
- Elementary and Secondary Elementary Act of 1965 (20 U.S.C. 6301 et seq.) as amended by the No Child Left Behind Act of 2001.
- Grissmer, D. (Ed.). (1999). Class size: Issues and new findings (Special issue). *Educational Evaluation and Policy Analysis*, 21 (2).
- Kaestle, C.F. (1993). The awful reputation of education research. *Educational Researcher*, 22 (1), pp. 26-31.
- Miller, D.W. (2001, July 13). The problem with studying vouchers. *The Chronicle of Higher Education*, pp. A14-A15.
- National Clearinghouse of Comprehensive School Reform. (undated) <http://www.csrrclearinghouse.org/>

- 
- National Education Association. (2002). *Charting new frontiers: Creating high performing schools*. Washington, DC: NEA. [\\_\\_\\_\\_\\_](#)
- National Research Council. (2001). *Scientific inquiry in education*. Committee on Scientific Principles for Education Research. Shavelson, R.J. and Towne, L. (eds.) Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Northwest Regional Educational Laboratory. (2001). *The catalogue of school reform models*.. Portland, OR: NWREL.
- Peterson, P.E., Howell, W.G., and Greene, J.P. (1999). An evaluation of the Cleveland voucher program after two years. Program on Education Policy and Governance Research Paper. Kennedy School of Government, Harvard University. [Online]. Available: <http://www.ksg.harvard.edu/pepg/>. [2001, August 21]
- Putnam, R.D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6, pp. 65-78.
- Ritter, G.W., and Boruch, R.F. (1999). The political and institutional origins of a randomized controlled trial on elementary school class size" Tennessee's Project STAR. *EEPA*, 21 (2), pp. 111-125.
- Slavin, R.E., and Fashola, O.S. (1998). *Show me the evidence! Proven and promising programs for America's schools*. Thousand Oaks, CA: Corwin Press.
- Southwest Educational Development Laboratory (undated). *CSRD Awards Database*. (<http://www.sedl.org/csrd/awards.html>)
- Sroufe, G.E. (1997). Improving the "awful reputation" of educational research. *Educational Researcher*, 26 (7), pp. 26-28.
- Willinsky, J. (2001). The strategic education research program and the public value of research. *Educational Researcher*, 30 (1), pp. 5-14.

**APPENDIX 1**  
**RESOURCES FOR SCHOOLS**

Type of Information	Where Can You Find this Information?
Research abstracts	<ul style="list-style-type: none"> <li>▪ ERIC Clearinghouse; <a href="http://ericir.syr.edu/Edu/">ericir.syr.edu/Edu/</a></li> <li>▪ National Clearinghouse for Comprehensive School Reform; <a href="http://www.goodschools.gwu.edu/csrl/ld.htm">www.goodschools.gwu.edu/csrl/ld.htm</a> , click on “CSR Library”</li> </ul>
Description of theoretical base (the idea behind the reform and why it should work).	<ul style="list-style-type: none"> <li>▪ Materials from model or program developers</li> <li>▪ <i>Catalog of School Reform Models</i>; NWREL (2001); <a href="http://www.nwrel.org/scpd/catalog/">www.nwrel.org/scpd/catalog/</a> and National Clearinghouse for Comprehensive School Reform <a href="http://www.goodschools.gwu.edu">www.goodschools.gwu.edu</a></li> <li>▪ <i>The Educators’ Guide to Schoolwide Reform</i> (1999). Herman, R., et.al. American Institutes for Research</li> <li>▪ National Clearinghouse for Comprehensive School Reform <a href="http://www.goodschools.gwu.edu">www.goodschools.gwu.edu</a></li> </ul>
Evidence of effectiveness of an instructional practice	<p><i>Individual studies:</i></p> <ul style="list-style-type: none"> <li>▪ Program or model developers</li> <li>▪ Education journals (e.g., American Educational Research Journal, American Journal of Education, Educational Leadership, Educational Research and Evaluation, Educational Researcher, Harvard Educational Review, Journal of Education for Students Placed at Risk, Review of Educational Research, School Effectiveness and School Improvement)</li> </ul> <p><i>Research summaries:</i></p> <ul style="list-style-type: none"> <li>▪ Slavin, R.E. and Fashola, O.S. (1998). <i>Show me the evidence! Proven and promising programs for America’s Schools</i>. Thousand Oaks, CA: Corwin Press.</li> <li>▪ Education journals (e.g., Review of Educational Research)</li> </ul>
Evidence of effectiveness of a set of practices or a program	<ul style="list-style-type: none"> <li>▪ Program or model developers</li> <li>▪ Education journals (e.g., American Educational Research Journal, American Journal of Education, Educational Leadership, Educational Research and Evaluation, Educational Researcher, Harvard Educational Review, Journal of Education for Students Placed at Risk, Review of Educational Research, School Effectiveness and School Improvement)</li> <li>▪ National Clearinghouse for Comprehensive School Reform <a href="http://www.goodschools.gwu.edu">www.goodschools.gwu.edu</a></li> <li>▪ <i>The Educators’ Guide to Schoolwide Reform</i> (1999). Herman, R., et.al. American Institutes for Research</li> <li>▪ Slavin and Fashola (1998) – see above</li> </ul>

Evidence of replicability	<ul style="list-style-type: none"> <li>▪ Program or model developers</li> <li>▪ Education journals (e.g., American Educational Research Journal, American Journal of Education, Educational Leadership, Educational Research and Evaluation, Educational Researcher, Harvard Educational Review, Journal of Education for Students Placed at Risk, Review of Educational Research, School Effectiveness and School Improvement)</li> <li>▪ The Educators’ Guide to Schoolwide Reform (1999). Herman, R., et. al. American Institutes for Research</li> <li>▪ Southwest Educational Development Laboratory (SEDL). CSRD Awards Database. <a href="http://www.sedl.org/csrd/awards.html">http://www.sedl.org/csrd/awards.html</a></li> </ul>
---------------------------	--

## APPENDIX 2

Definition of Scientifically based research (ESEA; Title IX, Part A:

(37) Scientifically based research –“The term ‘scientifically based research’—

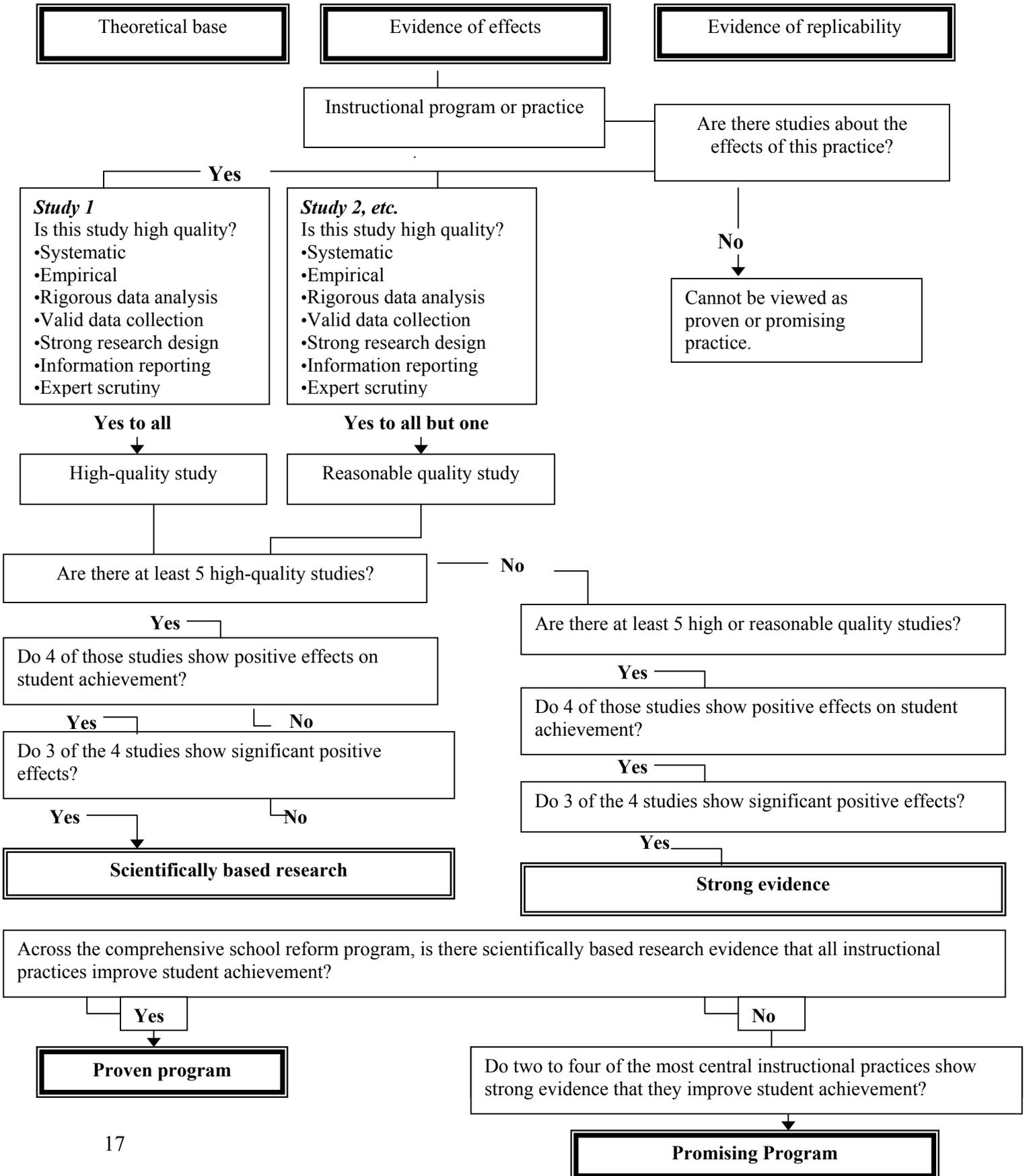
(A) means research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs; and

(B) includes research that—

- (i) employs systematic, empirical methods that draw on observation or experiment;
- (ii) involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;
- (iii) relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators;
- (iv) is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls;
- (v) ensures that experimental studies are presented in sufficient detail and clarity to allow for replication or, at a minimum, offer the opportunity to build systematically on their findings; and
- (vi) has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review.”

APPENDIX 3

Quality of Research Decision Tree



---

This booklet was designed to help school staffs and technical assistance providers become critical consumers of educational research and critically review the evidence of effects on student achievement. The information in the text is provided graphically on page seventeen in the form of a “decision tree”. The following example illustrates how the decision tree might be used.

Example:

Elementary School A is determined to take a comprehensive approach to reforming their school. Using the eleven CSR components, they have created a program that will affect all students and all teachers. Their program includes an externally developed model that emphasizes developing a family-style learning environment to help develop higher order thinking skills for all students, with a special focus on instruction. The CSR model has been used on 600 schools across the country, and the model developers provide professional development and materials.

The school will supplement that model with another change in instruction (cross-age lessons) and with increased parent involvement (by having parents act as tutors using the school’s curriculum). Both of these practices are consistent with the central idea of developing students’ higher order thinking skills in a family-style learning environment.

Before finalizing this plan, the staff wants to make sure that there is research evidence supporting the practices and programs they are considering. First, they look at the theory behind the externally developed model, asking questions such as, “How does a family-style learning environment work?” “Why does a family-style learning environment improve higher order thinking skills?” and “What is involved in creating this environment?” In asking these questions, they are determining the theoretical basis of the model and ensuring that all of the school’s teachers have the same picture in their heads of what the ideal family-style learning environment looks like, and why.

Next, the school staff critically reviews the evidence of effects. They begin by identifying the central instructional practices of their proposed program, such as cross-age instruction. They look for research studies on that practice and any others they are considering. Let’s say the staff finds seven studies on cross-age instruction. They would then determine the quality of each study, using the criteria summarized in the decision tree and explained in the *Guidelines for Judging the Quality of a Study*. Of the seven studies on cross-age instruction, they might determine that four are of high quality, two of reasonable quality, and one is not very good at all. Since there are at least four high quality studies, the school staff would move to investigating the findings. They might discover that three of the studies showed significant results indicating that students’ reasoning skills improved after participating in cross-age instruction, with two showing that students had gained more than 10 percentile points on state tests, an educationally significant finding. Using this evidence, the staff would determine that the instructional practice of cross-age instruction is a proven practice. In the same manner, the school staff would review the other instructional practices embedded in their comprehensive program, such as small group project work.

If all of the instructional practices can be shown to be effective as in the example above, then the CSR program would be “proven.” If the core—but not all—instructional practices can be shown to be effective, then the CSR program would be “promising.”

Finally, the school staff would want to consider implementation and replicability. Although there are 600 schools using the CSR model that is part of this school’s program, no schools are known to be using their exact CSR program, and hence there is no evidence that this combination of components will work together. Therefore, school staff will rely especially on the research evidence that indicates that each of these components works well in isolation (evidence of effects) and that the different components seem consistent with a common theme (theoretical base).