

CSRQ Center Report on Elementary School Comprehensive School Reform Models

Executive Summary

CSRQ Center Report on Elementary School Comprehensive School Reform Models provides a scientifically based, consumer-friendly review of the effectiveness and quality of 22 widely adopted elementary school comprehensive school reform (CSR) models. Each model is profiled and rated in the following categories:

- Category 1: Evidence of positive effects on student achievement
- Category 2: Evidence of positive effects on additional outcomes
- Category 3: Evidence of positive effects on parent, family, and community involvement
- Category 4: Evidence of link between research and the model's design
- Category 5: Evidence of services and support to schools to enable successful implementation

The report provides education stakeholders with a decision-making tool to help them sort through options about the hundreds of elementary school improvement choices available to meet local needs. The reviews are intended to clarify options, not to point to or endorse “best buys” from among the 22 models reviewed. Together, the reviewed models represent a significant portion of the total number of schools implementing elementary school CSR models. Each model serves a minimum of 20 elementary schools in at least three states and is available for adoption in almost all states.



AMERICAN
INSTITUTES
FOR RESEARCH®

The Comprehensive School Reform Quality Center (CSRQ Center) is funded by the U.S. Department of Education's Office of Elementary and Secondary Education, through a Comprehensive School Reform Quality Initiative Grant (S222B030012), and is operated by the American Institutes for Research (AIR) (<http://www.air.org>). Since 1946, AIR—a nonprofit social science research and technical assistance organization—has engaged in a variety of research, evaluation, technical assistance, and communication projects that help to make research relevant to education policy-makers and practitioners. The views expressed in this report do not necessarily reflect those of the U.S. Department of Education, AIR, or the CSRQ Center.

For more information on the CSRQ Center or to download reports, visit our Web site (<http://www.csrq.org>), contact us via e-mail (csrqa@air.org), or call our toll-free telephone number (866-544-8686).

Background on CSR

For the past decade, attention has focused increasingly on various strategies to improve schools and student achievement. CSR is one approach. It has been tried in thousands of schools nationwide, mostly for high-poverty students in low-performing schools. This trend is driven by the recognition that school improvement efforts are complex and require a coordinated, systematic approach that addresses every aspect of a school—including curriculum, instruction, governance, scheduling, professional development, assessment, and parent, family, and community involvement. Rather than use individual, piecemeal programs or approaches, effective CSR is meant to integrate research-based practices into a unified effort to raise student achievement and achieve other important outcomes, such as reducing dropout rates or improving behavior.

Many schools that adopt a CSR approach choose an external model to provide a research-based, replicable set of practices. These CSR models, or service providers, offer “blueprints” to help a school make improvements in a number of areas. Ideally, all models are designed based on research, and although they vary in focus, philosophy, and method, all help a school to raise student achievement. To support implementation, CSR models typically provide schools with materials, professional development, and hands-on assistance.

The CSR approach gained momentum with the 1997 passage of the federal Comprehensive School Reform Demonstration Program. Through this program, Congress provided dedicated funding to support the adoption of CSR strategies throughout the country. The 2001 Elementary and Secondary Education Act, also known as the No Child Left Behind (NCLB) Act, gave further impetus to the CSR approach by changing it from a demonstration project into a full-fledged federal program called the Comprehensive School Reform Program. According to the NCLB Act, CSR models must be scientifically based. This means that a model or approach must demonstrate strong research evidence that it can improve students’ academic achievement. Today, regardless of the funding source, the use of CSR is likely to remain an important strategy for improving schools, particularly those that fail to make Adequate Yearly Progress.

Since the mid-1990s, approximately 6,000 schools, serving several millions of students, have used federal funds to adopt more than 500 distinct CSR models and approaches. So far, overall results of the CSR approach have demonstrated promise, with some models helping schools make significant student achievement gains. For example, a 2002 systematic analysis by Dr. Geoffrey Borman and his colleagues of the student achievement outcomes of 29 leading K–12 CSR models reported that “the overall effects of CSR are statistically significant, meaningful, and appear to be greater than the effects of other interventions that have been designed to serve similar purposes and student and school populations” (Borman, Hewes, Overman, & Brown, 2002, p. 33).

CSR models, such as those described in this report, are promising because they are research based and provide the training and other supports needed to encourage a coordinated approach to achieve student success. However, the research evidence indicates that some programs are more effective than others and that their results vary greatly—even with the effective models—depending on the quality of implementation.

A Need for Consumer Information

To date, education stakeholders at the national, state, and local levels have had few objective, rigorous, and consumer-friendly sources to turn to when making choices from among the hundreds of CSR models available for adoption.

This report is intended to serve as a consumer guide that helps decision makers sort through claims about which approaches could truly meet the needs of students. It is the most extensive and comprehensive review of elementary school CSR models ever issued. To prepare this report, the Comprehensive School Reform Quality (CSRQ) Center reviewed nearly 800 existing studies on 22 widely implemented elementary school CSR models. To conduct our analysis, we used rigorous standards that are aligned with the requirements for scientifically based research established by NCLB.

Sorting through and making sense of claims made by researchers and model developers is hard work, even for research scientists with years of training and experience. Despite substantial advances in developing standards and processes for judging and “adding up” the evidence in education, areas of disagreement exist in the research community. Even when the procedures necessary for reviewing and comparing large numbers of studies exist, the process is often complex and painstaking. Therefore, education decision makers often turn to others, such as the CSRQ Center, to sort through the evidence and report it as “actionable” information.

The CSRQ Center Rating System

The production of this report was guided by the CSRQ Center’s Quality Review Tool (QRT). The QRT provides the criteria and procedures for independent, fair, and credible model reviews. The QRT development process involved several steps. First, CSRQ Center staff developed review frameworks in consultation with some of the nation’s most respected education researchers, program evaluators, and school improvement experts. The QRT also drew on prior and current efforts to conduct rigorous research reviews, including standards set by the U.S. Department of Education’s What Works Clearinghouse. Then, the QRT was reviewed and revised with the help of the CSRQ Center’s Advisory Committee, a nationally respected panel of experts that includes leading education practitioners, methodologists, and researchers from a variety of fields, including education, sociology, psychology, and economics.

Using the QRT, the CSRQ Center applied separate rubrics for each of the five categories of effectiveness and quality (listed previously) to arrive at its ratings, which are expressed by a set of symbols. In general, the rubrics resulted in the following ratings:

- **Very strong** rating is symbolized by a fully shaded circle (●). This is the “highest” rating provided by the CSRQ Center.
- **Moderately strong** rating is symbolized by a three-fourths shaded circle (◐). This is the “second highest” rating.
- **Moderate** rating is symbolized by a half-shaded circle (◑). Models receiving this rating may still have notable evidence of positive outcomes but not as strong as those receiving the ratings above.
- **Limited** rating is symbolized by a one-fourth shaded circle (◒). This rating indicates that while the CSRQ Center found some evidence of effectiveness, more rigorous research and evidence needs to be conducted on the model to fully support its effectiveness or quality on the category reviewed.
- **Zero** rating is symbolized by a circle with a horizontal slash (⊘). This rating means that while we found evidence that allowed the CSRQ Center to provide a rating for a category or subcategory, none of the studies were of sufficient quality to be counted as reliable evidence.

- **Negative** rating is symbolized by a circle with a minus sign (⊖). This rating indicates that we found strong evidence of detrimental effects in a given category or subcategory. In practice, we did not find any evidence of this kind for any model.
- **No Rating** is symbolized by “NR” in a circle (Ⓝ). This rating indicates that the model has no studies (i.e., evidence) available for review in a category or subcategory.

CSRQ Center Findings

Our report’s overall findings on Category 1 (Evidence of Positive Effects on Student Achievement) are similar to those of prior studies on CSR; that is, that models vary widely in both the number of rigorous studies and evidence that support their claims and in their effectiveness and quality when compared to each other. Our rating process for Category 1, and Categories 2 and 3, is complex and combines two elements to provide a single rating:

- The strength of the evidence based on the causal validity of the research design (e.g., how reliable and credible is it?)
- The strength of the reported impact or effect (e.g., does the model raise student achievement a little or a lot?)

For more than one third of the models, the CSRQ Center was able to identify only 10 or fewer studies that seemed to be relevant for our review of their overall evidence of positive effects on student achievement. In contrast, one model (Direct Instruction) had more than 50 and another (Success for All) had more than 100 studies that were originally considered for review in this category. In Category 1, after screening nearly 800 studies for quality, we found about 80 that met CSRQ Center standards. Again, these were unevenly distributed, with nearly one fourth of the models having no studies that met CSRQ Center standards and with five models (America’s Choice School Design, Direct Instruction, Literacy Collaborative, School Development Program, and Success for All) having five or more studies that met CSRQ Center standards. Table 1 summarizes the quantitative study findings that were used to rate evidence of overall positive effects on student achievement. For Category 1 (Evidence of Positive Effects on Student Achievement), we rated

- Two models as Moderately Strong (Direct Instruction, Success for All),
- Five models as Moderate (Accelerated Schools, America’s Choice School Design, Core Knowledge, School Renaissance, and School Development Program),
- Eight models as Limited (ATLAS Communities, Co-nect, Different Ways of Knowing, Integrated Thematic Instruction, Literacy Collaborative, National Writing Project, Modern Red SchoolHouse, and Ventures Initiative and Focus System), and
- Seven models as Zero (Breakthrough to Literacy, Comprehensive Early Literacy Learning, Community for Learning, Coalition of Essential Schools, Expeditionary Learning, First Steps, and Onward to Excellence II).

In reviewing findings for Category 1, readers should keep in mind that almost all of the models in the report serve high-poverty students in low-performing schools. Thus, the evidence of effectiveness they present is for success in educating students in highly challenging conditions.

The research base on which to rate models in Categories 2 and 3 is relatively sparse. It is important to note that a rating of Limited or higher in these categories indicates that the research on a model provides evidence of *positive impact* on additional outcomes for students, teachers, schools, family, and communities. Few of the models reviewed by the CSRQ Center had evidence that met CSRQ Center standards in these categories. The models that reported evidence of additional outcomes that met CSRQ Center standards in this category, which provides consumers with more information, are to be commended. All models are encouraged to seek and present this information in future evaluation reports.

The rating system for Categories 4 and 5 depended on several elements: (a) evidence of link between research and the model's design, (b) evidence that the model provider offers services and support to enable successful implementation, and (c) evidence that the model provider offers professional development and technical assistance to enable successful implementation. The same rating scale and symbols were used to rate Categories 4 and 5 as were used to rate Categories 1–3; however, the meanings of the ratings are category specific.

For Categories 4 and 5, the majority of the models reviewed provided moderate to strong evidence that they can provide a link between research and the model's design. The majority of the models reviewed also provided equally strong evidence that they can provide services and supports needed by schools to enable successful implementation. Given the importance of implementation to the success of any schoolwide reform, consumers who select models that have “lower” rankings in evidence of effects on student outcomes may still experience success if they implement the models faithfully. Table 2 summarizes the basic model information and model ratings for Categories 1–5.

Conclusions

When using the findings presented in this report to guide decision making, readers should keep in mind the following observations:

- **Although no model reviewed received a Very Strong rating for Category 1, several had a solid body of evidence.** These models may have received “higher” ratings if their outcomes had been more consistently positive or been supported by a few more studies that met CSRQ Center standards.
- **CSRQ Center standards were deliberately set very high.** We felt that, in an environment that requires scientifically based research for many program adoption decisions and features rigorous review standards set by the What Works Clearinghouse, it was important to give consumers confidence that models in this report were reviewed against stringent standards. Given the ongoing research that many of the models reviewed are currently conducting, we expect that our updated report, to be published in fall 2006, may result in “higher” ratings for several models.
- **A “lower” rating does not necessarily mean that a model is less effective than one with a “higher” rating.** It may mean that its effectiveness has not yet been established through rigorous research. All model providers should continue to conduct research to ensure consumers that (a) the model *may* work under *some* conditions and (b) the model *does* work under *most* conditions.
- **The ratings in this report are for a model's evidence at the elementary school level.** Some of the models reviewed in this report also provide services at the secondary school level. We did not review the studies

reporting outcomes at those levels to derive our ratings for this report. Therefore, it's quite possible that some models may have stronger evidence of effectiveness and quality when taking into account outcomes in middle and high schools. The CSRQ Center will issue a similar report for CSR secondary school models in fall 2006.

- **The purpose in providing ratings is to clarify options for decision makers, not to pick “winners” and “losers.”** Models that received “lower” ratings in Category 1 may have “higher” ratings in other categories. Education consumers should take a holistic view of the evidence presented across all five categories to make the best decision to meet locally defined needs.

The education community increasingly turns to research to help sort through its school improvement options. This reliance on research helps to satisfy the NCLB requirement that school improvement efforts be driven by scientifically based research. More importantly, it helps to meet the urgently felt need on the part of educators and policymakers to ensure that their efforts improve the lives of children. This report is issued by the CSRQ Center in the hopes that the information and analysis it contains makes a further contribution to making research relevant to improving the education of our nation's students.

Reference

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2002). *Comprehensive school reform and student achievement: A meta-analysis*. Baltimore: Center for Research on the Education of Students Placed at Risk, Johns Hopkins University.

Table 1. Quantitative Study Findings Used to Rate Evidence of Overall Positive Effects on Student Achievement

Comprehensive School Reform Model	Number of Studies					Number of Findings	Percentage of Positive Findings
	Initially Relevant	Eligible for Full Review	Meeting Standards	Conclusive	Suggestive		
Accelerated Schools PLUS	37	7	3	3	0	9	33.3
America's Choice School Design	16	7	7	6	1	16	47.9
ATLAS Communities	13	2	1	1	0	5	20.0
Breakthrough to Literacy	10	6	0	0	0	0	N/A
Comprehensive Early Literacy Learning	1	0	0	0	0	0	N/A
Community for Learning	6	0	0	0	0	0	N/A
Coalition of Essential Schools	11	1	0	0	0	0	N/A
Co-nect	22	5	2	2	0	9	22.2
Core Knowledge	20	5	3	3	0	12	50.0
Different Ways of Knowing	8	2	1	0	1	5	100.0
Direct Instruction (Full Immersion Model)	56	24	12	10	2	21	57.1
Expeditionary Learning	24	4	1	0	1	1	0.0
First Steps	3	0	0	0	0	0	N/A
Integrated Thematic Instruction	11	2	2	1	1	19	73.7
Literacy Collaborative	17	6	6	0	6	26	50.0
Modern Red SchoolHouse	23	3	2	0	2	4	100.0
National Writing Project	7	2	1	1	0	1	100.0
Onward to Excellence II	8	1	1	0	1	2	0.0
School Development Program	9	2	2	1	1	3	100.0
School Renaissance	31	8	5	3	2	12	50.0
Success for All	115	55	33	31	2	91	52.2
Ventures Initiative and Focus System	3	2	1	0	1	2	100.0
TOTAL	451	144	83	62	21	238	

Note. Some studies examined more than one comprehensive school reform model. Such studies were reviewed and counted once per model discussed. In these cases, a single study is counted more than once, and therefore, the total of the columns in this table exceeds the total number of studies reviewed.

Key:

Initially Relevant: Of the nearly 800 studies screened, the number of studies per model found to be relevant to this review.

Eligible for Full Review: The number of studies per model that used research designs that were sufficiently rigorous and included student achievement outcomes.

Meeting Standards: The number of studies per model considered *suggestive* or *conclusive* according to CSRQ QRT causal validity rubrics.

Conclusive: The number of studies per model that used a rigorous research design (e.g., experimental, quasi-experimental) with no critical threats to validity.

Suggestive: The number of studies per model that used a less rigorous research design (e.g., longitudinal) with no critical threats to validity.

Number of Findings: The total number of individual measured outcomes found in the studies that met standards.

Percentage of Positive Findings: The percentage of total findings in the studies that met standards that were statistically significant and indicated that a model had a positive impact. The N/A designation provided in this column indicates models in which zero studies met CSRQ Center standards.

Table 2. Summary of Basic Information by Model

Comprehensive School Reform Model	Grade Levels Served	Number of Schools	Year Introduced in Schools	Costs (Year 1)	Evidence of Positive Overall Effects	Evidence of Positive Effects for Diverse Student Populations	Evidence of Positive Effects in Subject Areas	Evidence of Positive Effects on Additional Outcomes	Evidence of Positive Effects on Parent, Family, and Community Involvement	Evidence of Link Between Research and the Model's Design	Evidence of Readiness for Successful Implementation	Evidence of Professional Development/ Technical Assistance for Successful Implementation
Accelerated Schools PLUS	K–8	143	1986	\$61,500			Reading and Math:					
America's Choice School Design	K–12	349	1998	\$70,000–\$105,000			Reading, Math, and Writing:					
ATLAS Communities	K–12	100	1993	\$60,000–\$80,000			Math:					
Breakthrough to Literacy	Pre-K–3	1,924	1992	\$15,500–\$17,500/classroom								
Coalition of Essential Schools	K–12	600	1984	Varies								
Community for Learning	K–12	150	1990	\$35,100								
Comprehensive Early Literacy Learning	K–12	812	1994	\$50,000								
Co-nect	K–8	200	1993	\$68,000/500 students			Reading, and Math:					
Core Knowledge	K–8	534	1990	Varies			Reading: Math, Science, and Social Studies:					
Different Ways of Knowing	K–12	500	1989	\$70,000			Reading, Math, Science, and Social Studies:					
Direct Instruction (Full Immersion Model)	K–8	64	1968	\$74,500			Reading: Math:					

Table 2. Summary of Basic Information by Model (continued)

Comprehensive School Reform Model	Grade Levels Served	Number of Schools	Year Introduced in Schools	Costs (Year 1)	Evidence of Positive Overall Effects	Evidence of Positive Effects for Diverse Student Populations	Evidence of Positive Effects in Subject Areas	Evidence of Positive Effects on Additional Outcomes	Evidence of Positive Effects on Parent, Family, and Community Involvement	Evidence of Link Between Research and the Model's Design	Evidence of Readiness for Successful Implementation	Evidence of Professional Development/ Technical Assistance for Successful Implementation
Expeditionary Learning	K–12	150	1993	\$55,000–\$65,000								
First Steps	K–8	335	1989	N/A								
Integrated Thematic Instruction	K–12	27	1984	\$76,500			Reading: Math and Science:					
Literacy Collaborative	K–6	584	1993	\$60,000			Reading:					
Modern Red SchoolHouse	K–12	300	1996	\$50,000–\$100,000			Reading, and Math:					
National Writing Project	K–12	189 university project sites	1974	N/A			Writing:					
Onward to Excellence II	K–12	1,000+	1981	\$18,000								
School Development Program	K–12	131	1968	Varies			Reading, and Math:					
School Renaissance	Pre-K–12	189	1986	\$56,884			Reading: Math:					
Success for All	K–8	1,400+	1987	\$88,580			Reading: Math, Science, and Social Studies:					
Ventures Initiative and Focus System	K–12	32	1981	Varies			Reading and Math:					

Note. Readers are encouraged to use this table in conjunction with the entire report, which explains in detail how the approaches were reviewed and rated. The report also provides detailed information about each model's ratings and offers in-depth descriptions of each model's services.

Table 2. Summary of Basic Information by Model (continued)

Key:

Grade Levels Served: Although this report focuses on a review of models implemented at the elementary school level, the grade levels served represents the full range of grades that the model serves.

Number of Schools: This reflects the number of schools using the model as reported by the model provider. This number includes all schools regardless of the length of time implemented or the level of implementation.

Year Introduced in Schools: This date refers to the year in which schools first implemented the model. This is included so that readers can judge whether the ratings are influenced by the relative newness of the model.

Costs (Year 1): The costs are estimates provided by the model provider. The full report provides additional details on costs for each model.

Evidence of Positive Overall Effects: This rating focuses on a model's overall effects on student achievement. The rating is a function of the number of studies that were rated as *suggestive* and *conclusive*, the percentage of findings in the suggestive and conclusive studies that demonstrated a positive impact, and the average effect size of those findings. The final rating reflects the amount of rigorous research and the strength of the effects reported in that research. The full report provides complete information about the methodology used to produce all ratings in this report.

Evidence of Positive Effects for Diverse Student Populations: This rating refers to positive effects for the achievement of students from diverse backgrounds, such as low socioeconomic status, minority, special needs, or English language learners.

Evidence of Positive Effects in Subject Areas: This rating refers to positive effects on achievement in specific subject areas, such as reading, math, writing, science, or social studies.

Evidence of Positive Effects on Additional Outcomes: This rating refers to positive effects on additional outcomes, such as student discipline, student attendance, school climate, retention/promotion rates, or teacher satisfaction.

Evidence of Positive Effects on Parent, Family, and Community Involvement: This rating refers to positive effects for improvement in family and community involvement, such as involvement in school governance, participation in family nights, or homework support.

Evidence of Link Between Research and the Model's Design: This rating refers to evidence that the model developer can provide explicit links between research and the core components of the model. Core components are considered essential to successful implementation.

Evidence of Readiness for Successful Implementation: This rating refers to evidence that the model provider ensures initial commitment from schools, tracks and supports full implementation, and helps schools allocate resources for successful implementation.

Evidence of Professional Development/Technical Assistance for Successful Implementation: This rating refers to evidence that the model provider offers comprehensive training opportunities and supporting materials, ensures that professional development effectively supports full model implementation, and develops the school's internal capacity to provide professional development.

 = Very Strong	 = Moderately Strong	 = Moderate	 = Limited	 = Zero	 = Negative	 = No Rating
---	---	--	---	--	--	---