

**CSRQ Center Report on Middle and High School CSR Models:
Frequently Asked Questions
October 2006**

The following is set of answers to frequently asked questions that the CSRQ Center often receives regarding our CSRQ Center Reports series (<http://www.csrq.org/reports.asp>). We have adapted them to answer specific questions that readers may have regarding our latest report on the evidence of effectiveness and quality of 18 widely adopted middle and high school improvement models. However, most answers apply equally well to our two previous reports, on 22 elementary school CSR models and 7 Education Service Providers. Altogether, the Center has now rated nearly 50 of the most widely adopted whole school improvement models in the country. If you have further questions, please contact us at csrqa@air.org.

- **What is Comprehensive School Reform (CSR)?**

For the past decade, attention has focused increasingly on various strategies to improve schools and student achievement. One approach, comprehensive school reform (CSR), has been tried in thousands of schools nationwide, most of which are high poverty and low performing. 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 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 reduced dropout rates or improved behavior.

Currently, there are several hundred external CSR models (also known as service providers) that districts and schools can turn to support local comprehensive school reform efforts. The U.S. Department of Education provides background information on comprehensive school reform at <http://www.ed.gov/programs/compreform/index.html>. Additional guidance on schoolwide reform is provided by the Department in its Guidance on Designing Schoolwide Programs (<http://www.ed.gov/policy/elsec/guid/designingswpguid.doc>).

- **What are CSRQ Center reports?**

CSRQ Center reports provide consumer friendly reviews and guidance on the effectiveness and quality of widely adopted comprehensive school reform (CSR) models. Although there are hundreds of CSR models offering services, in practice there are less than 100 that are in wide usage throughout the country. The reports are designed to give education consumers the in-depth information they need to make the best possible adoption decisions to meet locally defined needs. Each report provides basic information on the CSR model, including the model's mission and focus, year introduced in schools, grade levels served, number of schools served, and costs as well as ratings on five indicators of quality and effectiveness.

Our reports provide education stakeholders a decision making tool to help them sort through their options regarding the numerous middle and high school improvement choices available to meet local needs. The reviews are intended to clarify options, not to point to or endorse "best buys." The CSRQ Center's report on middle and high school CSR models complements our *Works in Progress* report on secondary school improvement programs

issued in January 2005 (<http://www.csrq.org/reports.asp>), by providing a consumer guide to widely adopted CSR models serving middle and high schools.

- **How was this report produced?**

The CSRQ Center does not conduct field research. That is, we do not go out and do original research on the models reviewed in our reports. Rather we gather and review existing studies and evidence on the effectiveness and quality of CSR models to provide its ratings. The production of this report was guided by the CSRQ Center's Quality Review Tool, or QRT, which 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, piloted, and revised with the support 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.

- **What types of evidence do you rate?**

For each model, the report evaluates the following dimensions of quality:

- Category 1: Evidence of positive effects on student achievement
- Category 2: Evidence of positive effects on additional outcomes, such as reduced dropout rates or improved discipline.
- 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

Whenever possible, the CSRQ Center Reports offer information on model results for specific student groups or specific types of school settings.

- **How does your rating system work?**

Our rating process is complex and although the same ratings apply across all five categories, the information used to develop the ratings varies. For example, Categories 1-3 combine two elements to provide a single rating:

- The strength of the evidence based upon the causal validity of the research designs used to test the model's impact (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?).

Using the QRT (see above), the CSRQ Center applied separate rubrics for each category of effectiveness and quality listed above to arrive at its ratings, which are expressed by a common set of symbols. In general, the rubrics we used result 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 $\frac{3}{4}$ shaded circle. This is the “second highest” rating.
- **MODERATE** rating is symbolized by a $\frac{1}{2}$ 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 $\frac{1}{4}$ 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 this 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 studies that allowed the CSRQ Center to provide a rating for a category or subcategory, none were of sufficient rigor 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 sub-category.

- **What is reviewed in this CSRQ Center report?**

This *CSRQ Center Reports on Middle and High School CSR Models* provides a scientifically based, consumer-friendly review of the effectiveness and quality of 18 widely adopted middle and high school comprehensive school reform (CSR) models. Together, the reviewed models represent a significant portion of the total number of schools implementing middle and high school CSR models. Each model reviewed serves a minimum of 40 middle and high schools spread over at least 3 states, and is available for adoption throughout most of the country.

- **How did you select which CSR models to rate?**

The CSRQ Center followed a process through which we visited the Web sites and reviewed publicly available material on more than 30 school improvement models working at the middle and high school level to identify our ultimate sample of 18 CSR models. In the initial screening, we selected each CSR model based on: (1) whether it served a minimum of 40 middle and high schools in at least 3 states, and (2) was available for adoption in almost all states. This yielded 37 comprehensive school reform models. Then, the Center examined whether the CSR model’s design features met the 11 key CSR components identified by the U.S. Department of Education (Department of Education, 2002). The components used by the CSRQ Center for our comprehensiveness rating were governance, technical assistance, classroom practices, professional development, leadership development,

benchmarks/assessments, and curriculum. The final list of 18 models was chosen on the basis of its wide implementation (“market share”) and “comprehensiveness.”

- **How many studies did you look at to arrive at your findings?**

The CSRQ Center looked initially at approximately 1500 documents (studies, articles, and other materials that seemed like they might be relevant for our review), located through an extensive search process, to reduce the number to 197 studies that were considered for this review. Following the QRT process discussed above, the Center eventually identified 41 quantitative studies that could be analyzed to provide ratings in Categories 1-3. Categories 4 and 5 were rated by using other, mainly qualitative information, described in the report. Detailed information to answer this question is included in the “About this Report” and “Methodology” sections of the report.

- **Which CSR models rated best in your review?**

This CSRQ Center report **does not** rate “best buys” or promote the adoption of any particular model, even if it has a “very strong” rating. Our goal is to clarify options and provide the in-depth research review and information necessary to make decisions that meet local needs. Decision makers are urged to make “holistic” evaluations that include: (1) the information we present in the five categories of quality, (2) our in-depth profiles, and (3) a consideration of unique variables or characteristics of local settings.

As we point out in the report, educators or policymakers in one district may be more willing to take on the approach of one model over another, even if this means adopting an approach with less evidence of effectiveness. This is their choice. However, they should be aware that they are making a trade-off, and possibly taking a greater gamble on success if they do so. In addition, since implementation is so important to success, decision makers adopting “less proven” models must make a strong commitment to making sure the model works in their school or district. In order to help educators and policymakers make the best decisions possible, our reviews provide extensive background information describing each model.

We believe that schools and students will benefit greatly if decision makers read through these profiles carefully and then use evidence and their sound judgment to weigh their options.

- **What did your report find?**

The following are some of the most notable findings detailed in the report.

In terms of quantity of available evidence, four models stood out (America’s Choice, School Development Program, Success for All-Middle Grades, and Talent Development High School). In contrast, for over half of the models, the CSRQ Center was able to identify only 10 or less studies that seemed to be relevant for our review of their effectiveness in this category.

Overall, in Category 1: Evidence of Positive Effects on Student Achievement. In this critical category we rated:

- Four models as Moderate (America's Choice, School Development Program, Success for All-Middle Grades, and Talent Development High),
- Six models as Limited (Expeditionary Learning, First Things First, KIPP, Middle Start, More Effective Schools, and Project GRAD) and
- Eight models as Zero (Accelerated Schools-PLUS, ATLAS Communities, Coalition of Essential Schools, High Schools that Work, Making Middle Grades Work, Modern Red SchoolHouse, Onward to Excellence II, and Turning Points).

As readers review findings in Category 1 they 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.

Other findings are that a vast majority of the models reviewed provide moderate to very strong evidence that they can demonstrate a link between research and the model's design, and equally strong evidence that they provide the services and supports needed by schools to enable successful implementation. Given the importance of implementation to the success of any schoolwide reform, this means that consumers who select models that have "lower" rankings in evidence of effects on student outcomes *may* still experience success if they implement these models faithfully.

- **You didn't rank any models as having *Very Strong* evidence that they have positive effects on student achievement. Does this mean that CSR models represent a weak school improvement approach?**

This report does not call into question the effectiveness of comprehensive school reform (CSR). Further, the report does not represent a study of the CSR approach as a whole, but rather the effectiveness and quality of selected CSR models. Comprehensive school reform is one of the most widely studied improvement approaches in the history of education and has demonstrated promising evidence of effectiveness in various studies. 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—most of which operated in Title 1 schools—reported that "the overall effects of CSR are 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).

It is true that none of the models reviewed received a "Very Strong" rating for Category 1. To receive this rating, models must provide considerable evidence indicating positive outcomes based on rigorous research, primarily randomized controlled trials (RCTs, also referred to as experiments) or solid quasi-experimental studies. In general, few RCTs have yet to be conducted on education programs, and very few have been conducted on CSR models. Therefore, it was hard, in this round of reviews, for a model to receive a "Very Strong" rating. Readers should be aware that our 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, that it is important to give consumers confidence that models in this report were reviewed against stringent standards.

However, several models with a solid body of evidence might have received "higher" ratings, had their outcomes been more consistently positive, or been supported by a few

more studies that met our standards. Furthermore, numerous models are paying increased attention to conducting rigorous research to demonstrate their effectiveness.

- **Many of the models in your report are not rated on their evidence of raising student achievement for specific sub-populations such as low income or minority students. Does this mean that they are not effective when they work in Title I or similar schools?**

No. Readers should not necessarily judge a “No Rating” or a “lower” rating on the category *evidence of positive effects on diverse student populations* as evidence that the model cannot be effective in Title I schools, or other schools with similar student populations. In fact, federally funded CSR models on average serve school populations with a poverty rate of about 70%. Thus, readers may interpret our overall rating on the category of positive overall effects on student achievement as an indicator of their effectiveness in working in challenging settings, such as Title I schools. However, we found that only a few models broke out their results in Title I or other schools in such a way as to permit the Center to provide ratings on model effectiveness for specific subpopulations. Those models that do provide this information are to be commended for providing additional consumer information. All models are encouraged to seek and present this information in future evaluation reports.

- **Why are CSRQ Center reports important to educators and policymakers?**

If, as a country, we are to realize the potential offered by comprehensive school reform and schoolwide improvement models, decision makers will need help in sorting through the diverse range of information and often competing claims made by researchers and providers about the effectiveness of various comprehensive school reform models. The CSRQ Center provides scientifically based, consumer-friendly reports on CSR model quality and effectiveness and the guidance needed to use this evidence to make effective, locally defined choices.

- **How can educators use this report?**

To date, educators have had few objective, rigorous, and consumer-friendly sources to turn to when making choices from among the hundreds of CSR models and improvement approaches available for adoption. This report is intended to provide a consumer guide that helps busy educators sort through claims about which approaches could truly meet the needs of students. The CSRQ Center, as a support tool for educators, has sorted through the evidence and reported on models in the form of reviews. Educators, individually or as school improvement teams, are encouraged to use the individual reviews, or the CSR database provided on the CSRQ Center’s Web site to compare and weigh their options. A solid adoption decision is only one step in ensuring effective improvement. Thus, educators are encouraged to conduct school needs assessments prior to investigating their adoption decisions. Once adoption decisions are made, educators must be willing to commit the time and effort to fully implement the improvement approach they have selected.

Educators are encouraged to use systematic approaches for implementing school reform models, such as the process described in the CSRQ Center’s *Moving Forward: A Guide for Implementing Comprehensive School Reform & Improvement Strategies*. Available free through our Web site, *Moving Forward* includes tools to assist educators in implementing comprehensive school reform models.

In addition to the CSRQ Center, the Center for Comprehensive School Reform and Improvement (www.csrclearinghouse.org) provides a variety of resources to support effective CSR—from judicious adoption to effective implementation.

- **How can policymakers use this report?**

It is particularly important in these times of increased accountability and scarce resources for policymakers to encourage the use of school improvement approaches that have demonstrated effectiveness and quality. The CSRQ Center’s reports are not intended to encourage the production of “lists” of approved models, but rather to encourage the use of high quality evidence when making model adoption decisions. Policymakers can contribute to building evidence-based practice in education by insisting that school improvement approaches that are selected are based on solid evidence, and then by providing the support necessary to ensure their full implementation. Research consistently demonstrates that even the most effective improvement approaches may take several years before they demonstrate results, and that effective implementation is the key to getting good results. Policymakers can help schools and districts choose strong approaches and then give them the time and support they need to get results.

- **How did the models included in the report participate in your review process?**

The CSRQ Center contacted each model at several points in the review process. These contacts were used to verify information that we had gathered from other public sources to ensure that our reviews were as accurate as possible. Our report is absolutely independent of any provider’s views.

As described in the “About this Report” and “Methodology” sections of the report, the CSRQ Center involved the reviewed models at several steps in the production of the report. After developing our own background information and profile of the model based on publicly available sources, we contacted models to inform them of the review and to speak to them informally to verify the information we had compiled. We held conversations with almost all providers. These conversations helped us to identify additional materials to collect, research to review, and individuals to contact for our report.

At the end of the process the CSRQ Center provided all models a draft of their review, and background information on the report so that they could comment on its accuracy. Each model was asked to review the draft and note if the review contained any inaccuracies. The CSRQ Center invited providers to share questions and concerns, and provide documentation for any information they felt needed to be corrected. In addition, the CSRQ Center encouraged interested providers to submit a two-page response letter to their review, to be published along with the report.

Many providers engaged in telephone and email communication with the CSRQ Center in the weeks following receipt of our draft review, and provided the CSRQ Center valuable insight and information to improve our report. The CSRQ Center considered all concerns or suggested edits for inclusion in the final narrative. The letters provided by the models give consumers additional information that they should weigh in making adoption decisions. All letters have been reproduced as submitted to the CSRQ Center and are available as part of our report.

- **How can I find out more about all the other CSR models that are not reviewed in your reports?**

Some education decision makers may be interested in CSR models, including new or smaller models that have not yet been reviewed by the CSRQ Center Reports. The Center provides a non-evaluative *CSR Model Registry* (<http://www.csrq.org/CSRProgramRegistry.asp>) on our Web site so that model developers have the opportunity to share their own information about models not included in the CSRQ Center Reports. Consumers can use this registry to further explore and compare their options.

- **The model that is used in my school is not included in your review. What does this mean? Is this a bad model?**

The fact that only 18 models are included in this review is not meant to indicate that these are the only CSR models worth considering for adoption. As a national “consumer reporting” organization the CSRQ Center tried to review as many widely adopted CSR models as our resources permitted. Given this limitation, our review covered only widely adopted CSR providers. However, the CSRQ Center provides a non-evaluative *CSR Model Registry* on our Web site (<http://www.csrq.org/CSRProgramRegistry.asp>) so that other model developers have the opportunity to share their own information about models not included in the CSRQ Center Reports. Consumers can use this registry to further explore and compare their options.

If the model used in your school was not reviewed, we encourage you to urge the model to submit information for the Registry. The information provided by the Registry closely parallels the rating categories we have established for our reviews. This will allow you, and others interested in the model you have adopted, to make some comparisons. Finally, readers should remember that a lack of rigorous evidence does not necessarily mean that a model is ineffective or of low quality. However, it does mean that it currently lacks sufficient evidence to support claims of effectiveness that all models, regardless of size, should provide in the future.

- **What is an effect size and why did you choose the range you use to rate model effectiveness?**

Effect sizes (ESs) are complex calculations, used to standardize measures of the impact of interventions. Generally, ESs are used to estimate gains (+) and losses (-) on achievement or other outcomes, where the differences between the groups being measured are expressed in standard deviations (SDs). ESs are used for two purposes in our report. First, they help us to describe results for individual studies. Second, we use them to calculate average ESs across various studies of the same model. These calculations help us to rate the effectiveness of individual models and to compare the impact of models against one another.

Based on a review of existing literature on ESs for CSR models and in consultation with experts, we set ranges in our ratings for moderate (+0.15 to +0.19), moderately strong (+0.20 to +0.24) and very strong (+0.25 and above) as components of our rating process. Because of differences among study designs and assessments, our determination of ESs can

only be considered a rough estimate of impact, allowing comparison among the various models.

- **Why didn't you look at studies conducted prior to 1980?**

When conducting systematic evidence reviews that require the identification, collection, and analysis of thousands of study findings, researchers must make important early decisions regarding where and how far back to look for evidence. For each CSR model, researchers searched educational databases (e.g. JSTOR, ERIC, EBSCO, Psycinfo, Sociofile, NWREL, DAI), web-based repositories (e.g. Google, Yahoo, Google Scholar), and two previous large-scale reports on comprehensive school reform (Herman et al., 1999; Borman et al., 2002).

In order to conduct as extensive a search as possible across multiple sources a limit had to be placed in how far back in time we searched for evidence. Based on a review of similar efforts being conducted by the What Works Clearinghouse and in consultation with our advisors, we chose 1980 as the earliest year for studies to qualify for our review. We felt that this was well within the "industry average" for this work, and provided nearly a quarter of a century for each model to demonstrate its effectiveness and quality. Few models have been in existence that long. Some models may have concerns that the 1980 limit excluded key findings on their model, which they consider definitive. In these instances, we encouraged the model providers to submit a letter to us citing these studies so that readers would be able to factor this evidence into their decisions.

- **How do your ratings compare to findings of prior similar studies?**

In format and approach, this study most closely resembles the Herman, et al, *An Educators' Guide to Schoolwide Reform* that was commissioned by leading national education organizations, and issued by the American Institutes for Research (AIR) in 1999. That report reviewed and rated the evidence of (a) positive effect on student achievement and (b) support that the reviewed developers provided to schools. The current report expanded the number of categories of quality and effectiveness reviewed to five, and created new processes and standards to conduct the review. Importantly, it "raised the bar" on the evidence needed to receive the "highest" rating in our Categories 1-3 (see description of categories above). We felt that our standards had to be set at a very high level, in an era that requires scientifically based research, and has the What Works Clearinghouse standards to serve as a guide.

In addition, while our study is not a meta-analysis, we drew important guidance from the systematic review of 29 CSR models produced by Dr. Geoffrey Borman and his colleagues in 2002. Their study served as a strong scientific anchor against which the CSRQ Center conducted its own rating work.

Direct comparisons of these prior studies with our review are difficult for a number of reasons. First, we focused on middle and high CSR models only while the other two looked at a set of K-12 models, only some of which serve middle and high schools. In addition the rules governing which quantitative studies to include differed in each review. As well, our rating framework and ranges differed, making it difficult to equate individual ratings from one study to the next. Finally, since time has elapsed the body of research evidence on models has grown, making it hard to predict the ratings the models we reviewed would have received using the alternative approaches taken by the other two studies.

However, we remind consumers that our report is not about ranking “winners” and “losers” but rather about providing solid information to promote evidence-based decision making. Too much of a focus on which model did “better” or “worse” in one review or another detracts from the effort to find the best models to meet local needs and conditions. In this regard we feel that this report builds on previous ones and provides education decision-makers the best evidence yet available to support informed choices.

- **Don't positive outcomes for any model depend on implementation?**

Absolutely. A consistent finding of research on school improvement programs and models—even the most effective ones—is that positive outcomes depend on effective implementation that is carried out by highly committed, adequately trained, and well supported educators.

Many of the models reviewed in this report have solid evidence that they provide the supports necessary for effective implementation. However, it is also true that research evidence indicates that—all other things being equal, including implementation—some programs have demonstrated greater effectiveness than others. In recognition of the fact that implementation can have a powerful impact on outcomes, the CSRQ Center excluded from our analysis of effectiveness those studies in which implementation level did not meet a minimal threshold level. In practice this is difficult since studies do not uniformly measure and present implementation information. To preserve transparency in our decision-making, we carefully documented all decisions regarding inclusion, exclusion, and ratings of studies based on this or any other factor.

Our report helps consumers identify which models currently present the strongest evidence of positive impacts on student achievement and other outcomes. Models presenting strong evidence will change over time, as they improve their operation, and as some demonstrate their effectiveness through rigorous research. In our report, consumers are offered information both about model effectiveness and about support for implementation, so that they can make a wise choice from among alternatives.