

PART A. JUSTIFICATION

A1. Circumstances Making Collection of Information Necessary

Comprehensive School Reform under No Child Left Behind Act of 2001, P.L. 107-110.

This request is for OMB approval of revised data collection associated with the *Longitudinal Assessment of the Comprehensive School Reform Program Implementation and Outcomes (LACIO)*. Sec. 1606 of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act (P.L. 107-110) mandates activities to be conducted by K-12 schools across the country under “comprehensive school reform,” and Sec. 1607 mandates the National Evaluation.¹

Along with the new authorization, Congress appropriated \$235 million for the Comprehensive School Reform program in fiscal year 2002.² This level of funding supported reform activities at an estimated 2,000 schools. The vast majority of these schools were Title I schools “in need of substantially improving” their student achievement levels.³

The federal funds were distributed on a formula basis to the states, who in turn made grants to districts to support the schools.

Each school received an average of over \$70,000 per year for three years. The modest award amounts and limited award durations were intended to signal a catalytic role for the federal funds - helping a school to initiate or advance its reform efforts - rather than serving as a long-term subsidy.

As of September 2005, the schools in the cohort studied through LACIO have completed the original funding cycle. In addition, current appropriation legislation will, at best, support the current CSR schools through their funding cycle. Consequently, this request for the revision of

¹ The new legislation continues comprehensive school reform initiatives started in FY1998 and originally established through appropriations, not authorizing legislation, the Appropriations Act for the U.S. Department of Education, P.L. 105-78.

² Total support for comprehensive school reform activities consisted of this appropriation plus another \$75 million from an account under the Fund for the Improvement of Education (FIE). Thus, the total amount of funds available to support the activities was actually \$310 million.

³ The earlier appropriations, (covering both comprehensive school reform and FIE allocations (see previous footnote) and number of schools supported are as follows: FY1998: \$145 million (about 1,800 schools); FY1999: \$145 million (to continue 1,800 schools along with about 441 new awards); FY2000: \$220 million (to continue about 2,241 schools and also to fund about 568 new awards); FY2001: \$260 million (to continue all the previous schools and fund about 139 new awards); and FY2002: \$310 million (to continue all the previous schools and fund an estimated 2,000 new awards).

OMB approval reflects a changed focus of the LACIO to address two concerns: (1) the extent to which the federal funds actually played the anticipated catalytic role and schools continue reform after CSR funding ended; and (2) the lessons from CSR that can be applied to all schoolwide Title I programs.

Evaluation Goals and Questions. The *No Child Left Behind* legislation stipulates two broad goals for the LACIO:

1. To evaluate the implementation and results achieved by schools after 3 years of implementing comprehensive school reforms; and
2. To assess the effectiveness of comprehensive school reform in schools with diverse characteristics.

The original U.S. Department of Education solicitation for the National Evaluation articulated these two goals in terms of three specific evaluation questions:

Evaluation Question No. 1: To what extent have CSR schools made progress on state assessments in comparison to gains for schools in their state with similar characteristics?

Evaluation Question No. 2: How effective are various school reform activities, especially in diverse settings, and to what extent can school progress be linked to CSR reforms?

Evaluation Question No. 3: How have district policies and state policies affected CSR program implementation and comprehensive school reform?

In 2005, ED revised the LACIO to collect follow-up information on the extent to which reforms were sustained and to identify approaches for improving Title I schoolwide programs. In addition, given the current focus on high school reform, ED determined that the LACIO could conduct additional analyses and on-site data collection about Title I secondary schools engaged in reform efforts. Consequently, in addition to the previous three evaluation questions, the revised LACIO will address two new questions:

Evaluation Question No. 4: What implications can be drawn from CSR implementation and outcomes for reform in Title I schoolwides?

Evaluation Question No. 5: How effective are various school reform activities in secondary schools, and to what extent can school progress be linked to comprehensive school reform?

Evaluation Design

The evaluation of CSR will continue to analyze student data and school reform at multiple points in time, comparing CSR program and non-CSR program schools. The design will use multiple methods of data collection and analysis to increase the rigor of the entire evaluation. The LACIO will continue to consist of three components (analyses of student achievement, surveys of school, district, and state reform activities, and field-based studies of reform activities), each contributing to answering the questions posed by ED.

For two of the components, the evaluation team will select a large sample of schools and a smaller sample of districts and states for an analysis of student achievement outcomes and a survey of CSR program-related activities. WestEd will also examine CSR-related activities at the district and state level. For the third component, a smaller sample of CSR program and non-CSR program schools will provide data for the field-based inquiry.

Analysis of student achievement

For all schools in the evaluation, the evaluation team will conduct a *quantitative analysis of changes in student achievement*. WestEd will collect the relevant student achievement data from databases maintained by the U.S. Department of Education. The analysis covers a random sample of 500 CSR program schools (representing about 20 percent of all schools that received new CSR program funds under P.L. 107-110) and 500 non-CSR program schools selected to match the CSR program schools demographically. This sample size will have sufficient power to detect differences with small to medium effect sizes. For each school, WestEd will analyze multiple test score points: one for the period prior to CSR program funding, three for the period of CSR program funding, and at least one for the period after federal program funding has ended. Earlier analyses examined changes in achievement scores over the first year of implementation for both CSR program and non-CSR program schools using multiple regression to control for school characteristics and baseline achievement. These analyses provided information on both achievement differences between the two groups of schools as well as the influence of CSR program status on achievement. The current phase of analysis will expand the knowledge about the relationship of CSR and achievement using multi-level modeling techniques to assess differences in achievement growth rates of CSR program and non-CSR program schools

Survey of school reform activities

The same sample of 1,000 schools will complete a *Survey of Reform Activities at CSR Program and non-CSR Program Schools* form to describe the various reform activities occurring

at the schools. The survey will examine the presence of the 11 components of Comprehensive School Reform included in the *No Child Left Behind Act of 2001*, and other elements (e.g., school organization) prior research shows to be associated with successful program implementation. WestEd will conduct two regression analyses. The first will focus on the reform components and state and district contexts associated with implementation of reform. The second will use student achievement as the dependent variable, with implementation (measured through survey responses) as the independent variable.

The strength of this method is that it can cover a large sample and the desired diversity of school conditions; the weakness is that it collects self-reported data (although the use of behaviorally anchored survey items that measure behaviors and are evidence-based, rather than attitudes and expectations, will reduce the error associated with response bias). Furthermore, to corroborate the survey data, as well as to provide a stronger evaluation, WestEd will use case studies of schools in their district and state contexts. The case studies will include classroom observations.

Field-based study of reform activities

Out of the larger sample of 1,000 schools, WestEd will select a subsample of 40 schools (half CSR program and half non-CSR program) to participate in the *Field-Based Study of Reform Activities at CSR Program and non-CSR Program Schools*. The original 15 CSR program and 15 non-CSR program schools will remain in the field-based study. An additional five schools will be secondary schools drawn from the larger CSR sample and matched with five comparable non-CSR high schools. The sample will have two foci: fieldwork in 10 of the 15 pairs of schools will focus on sustainability and integration with schoolwide Title I programs. The remaining five schools will constitute “best cases” among those found in the first three years of the evaluation. “Best cases” are schools that have high levels of CSR implementation **and** increased student achievement outcomes. If available, best cases will include schools that have successfully exited restructuring status under NCLB.

The component calls for two visits to each site, occurring during the first and second year of the revision of OMB approval. Each “site” consists of four entities:

1. A CSR-funded school;
2. A demographically matched non-CSR school (a school that has not received any federal CSR funds), located in the same district as the CSR-funded school;

3. The district within which the two schools are located; and
4. The state within which the district is located

By covering these four entities, the field-based component will address the new evaluation questions and also produce an understanding of the dynamic of the actual relationships among school, district, and state actions, policies, and practices.

During the site visits in the field-based component, the evaluation team will complete organizational protocols covering relevant events at the “site.” Data for the protocols will come from classroom observations, using a formal observation instrument, as well as other direct field observations, reviews of relevant school documents and materials, and discussions with school staff and parents.

Survey of district and state reform activities

The evaluation team will gather data from a sample of district officials as well as state officials at the states not included in the field-based data collection. Recent research shows that district and state actions, policies, and practices play an important role in school reform. These surveys, administered on-line, will expand on data collected during the site visits to document the conditions in districts and states that initiate, implement, and sustain reform activities in CSR schools and other Title I schoolwides.

The *district survey* will include a total of 65 school districts. The district official with supervisory responsibility for Title I schoolwides will complete the survey.

The *state survey* will include all 50 states. Respondents to the on-line survey will be the state CSR or Title I coordinator.

Scope of Data Collection.

The large sample of schools completing the *School Survey* will be a random sample of CSR-funded schools. It will be the basis for generalizing the findings of the LACIO to the entire universe - the estimated 2,000 schools that received funds for the 2002-2003 academic year. At the same time, the sample of 40 sites in the *field-based component* will come from the pool of 1000 schools in the large sample school component. This overlap, along with the similarity of the organizational survey instrument to be used in both components, will enable the evaluation team to use the site visit data to identify biases, if any, in the data from the self-reported survey.

The topics for the variables and data in the LACIO data collection derive directly from the provisions of P.L. 107-110, as well as guidance related to the Title I schoolwide program. The data in all of the LACIO’s three components will concentrate on determining whether schools

are sustaining CSR-related reforms, and will also cover school conditions that are related to school reforms beyond the CSR components.⁴ The data collection in all of the components, excluding the student achievement data from the existing U.S. Department of Education archive, is the subject of the present request for OMB clearance under the Paperwork Reduction Act.

Data Collection and Measures

Across the various data collection methods, the 11 components of the CSR program defined in P.L. 107-110 (see Exhibit 1) as well as the topics for Title I schoolwides guided the development of measures for all of the instruments. New data collection activities will begin in Spring 2006 and continue through the 2006-07 school year.

⁴ A more detailed discussion of the data collection activities is found later in section B2 of this OMB Clearance Request.

Exhibit 1
Eleven Components of Comprehensive School Reform
(Elementary and Secondary Education Act, as reauthorized by the *No Child Left Behind Act of 2001*)

SEC. 1606. Local Use of Funds

(a) Uses of Funds.

...to enable the schools to implement a comprehensive school reform program that –

- (1) employs 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;
- (2) integrates a comprehensive design for effective school functioning, including instruction, assessment, classroom management, professional development, parental involvement, and school management, that aligns the school's curriculum, technology, and professional development into a comprehensive school reform plan for schoolwide change designed to enable all students to meet challenging State content and student academic achievement standards and addresses needs identified through a school needs assessment;
- (3) provides high quality and continuous teacher and staff professional development;
- (4) includes measurable goals for student academic achievement and benchmarks for meeting such goals;
- (5) is supported by teachers, principals, administrators, school personnel staff, and other professional staff;
- (6) provides support for teachers, principals, administrators, and other school staff;
- (7) provides for the meaningful involvement of parents and the local community in planning, implementing, and evaluating school improvement activities consistent with section 1118 [see below];
- (8) uses high quality external technical support and assistance from an entity that has experience and expertise in schoolwide reform and improvement, which may include an institution of higher education;
- (9) includes a plan for the annual evaluation of the implementation of school reforms and the student results achieved;
- (10) identifies other resources, including Federal, State, local, and private resources, that shall be used to coordinate services that will support and sustain the comprehensive school reform effort; and
- (11) (A) 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
(B) has been found to have strong evidence that such program will significantly improve the academic achievement of participating children.

(b) Special Rule. A school that receives funds to develop a comprehensive school reform program shall not be limited to using nationally available approaches, but may develop the school's own comprehensive school reform program for schoolwide change as described in subsection (a).

SEC. 1607. Evaluation and Reports

(b) *Evaluation.* The national evaluation shall –

(1) evaluate the implementation and results achieved by schools after 3 years of implementing comprehensive school reforms; and

(2) assess the effectiveness of comprehensive school reforms in schools with diverse characteristics.

A2. Purpose and Uses of the Data

The U.S. Department of Education and other interested parties will use the data from the LACIO to assess the sustainability and student achievement outcomes from the comprehensive school reform provisions as stated in Sec. 1606 of the ESEA. It will also provide information that the U.S. Department of Education can use to strengthen schoolwide Title I programs by pointing to important processes and lessons learned from CSR. This information is intended to be useful to state and local school systems, including individual schools, in their efforts to achieve Adequate Yearly Progress (AYP) under the No Child Left Behind Act.

A3. Use of Technology to Reduce Burden

The evaluation team will use technology in a variety of ways, especially to reduce burden on the CSR schools. First, in the past few years WestEd amassed basic information about the schools on an electronic database created by the Southwest Educational Development Laboratory (SEDL), with support from the CSR program. The evaluation team will use this database to define the universe of CSR schools and to draw samples for the various LACIO components.

Second, the student achievement data for each school in the sample will be available through databases maintained by ED. These databases will be updated each year. The availability of such data will reduce the data collection burden placed on individual school sites. In cases where the database contains incomplete records, evaluators will contact the appropriate state and local officials for updated information.

Communication between the evaluation team and selected school, district, and state officials will occur through email, fax, and conference calls that take advantage of information technology and reduce burdens associated with paperwork. The communication will cover initial inquiries, the exchange of preliminary information, the scheduling and planning of site visits, and the review of draft reports. State and district officials will complete their surveys on line.

Throughout the evaluation, WestEd will provide a toll-free number and email addresses to respondents allowing them to contact the evaluation team with any questions or requests for

assistance. WestEd will print this information, along with the names of contact persons at WestEd and COSMOS on all data collection instruments.

A4. Efforts to Identify Duplication

The design of the LACIO is built upon the survey and research questions posed by previous research efforts, including the National Longitudinal Survey of Schools (NLSS), the Field-Focused Study, the National Study of Title I Schools (NSTS), and the Longitudinal Evaluation of the Effectiveness of School Interventions (LEESI). However, the LACIO is unique in that it combines elements of each study into a comprehensive data collection effort that allows for comparisons of CSR-program and non-CSR program schools and includes a small sample of schools that will participate in each data collection component. The variety of components within LACIO will allow researchers to make quantitative estimates of CSR-supported activities and efforts (especially among the 2002 cohort of schools) and describe how the implementation of such efforts is linked to education reform and student achievement.

WestEd will obtain the student achievement data for the sample of 500 CSR program and 500 non-CSR program schools from the ED databases, which contain outcome measures for schools in every state. Having access to these secondary data will allow researchers to reserve data collection efforts at the school, district, and state for only the most necessary data elements.

A5. Methods to Minimize Burden on Small Entities

The LACIO will collect data from few small entities, as most of the data sources will be school organizations (state, district, and local). The few small entities are likely to be associated with the external technical assistants and consultants who may be helping schools to implement comprehensive school reform. Only minimal information will be needed from these small entities, and so no significant impact on these data sources is expected.

A6. Consequences of Not Collecting the Data

The revision of LACIO will provide the U.S. Department of Education with a complete picture of the implementation and results achieved by schools after three years of implementing comprehensive school reforms and the extent to which both reforms and student achievement gains were sustained after funding ended. In addition, the evaluation will outline the effectiveness of comprehensive school reform in schools with diverse characteristics. Such answers are necessary to understand how federal and state funds serve as a stimulus for school reform.

The data collection efforts across the larger sample of CSR program and non-CSR program schools will allow researchers to comment upon the state of school reform across the nation. The LACIO's combination of student achievement data, descriptive survey data, and intensive field-based study comprises a design that builds on the strength of the combined methods.

A7. Special Circumstances

This information collection fully complies with 5 CFR 1320.5(d)(2).

A8. Federal Register Comments and Persons Consulted Outside the Agency

A notice about the study was published in the *Federal Register* on November 16, 2005.

The evaluation team will seek the expertise of persons outside the agency through the creation of a Technical Working Group (TWG). The TWG will advise the evaluation team on issues of school reform from the perspective of various stakeholders as well as methodological issues in evaluating CSR. We expect that the TWG will meet 7-8 times during the course of the study, with such meetings being tied to important events or tasks within the evaluation. Each TWG member will receive an honorarium of \$750 per day. The time commitment is relatively small, but the TWG will play an important role in providing insight and guidance in support of a successful evaluation. The TWG members are listed in Exhibit 2.

Exhibit 2

Members of the LACIO Technical Working Group

Member	Affiliation	Areas of Expertise
Geoffrey Borman	University of Wisconsin	Educational policy; comprehensive school reform; quantitative methods
Dan Bugler	Chicago Public Schools, Accountability, Research and Evaluation	District policy; research design
Kathleen de Marrais	University of Georgia	Qualitative methods
Elizabeth Hale	Institute for Educational Leadership	Educational policy
Deborah Hoffman	Principal, Madison WI	School policy
Valerie Lee	University of Michigan	Quantitative methods; school reform
Judith Preissle	University of Georgia	Qualitative methods
Sheila Rosenblum	Rosenblum-Brigham Associates	Institutional change
Sam Stringfield	University of Louisville	School improvement;

		comprehensive school reform
Malik Stewart	Delaware Department of Education	State policy; school improvement

A9. Payment or Gifts

The enormous pressures on school systems, in part due to increased assessment and accountability requirements, lead to their assigning a lower priority for participating in data collection efforts such as the LACIO. To indicate the importance of the work of the evaluation for informing federal, state, and district policies and practices on comprehensive school reform, schools participating in the LACIO will receive a special monetary payment. Past research shows such payments are a cost-effective way of increasing participation rates substantially (e.g., Dillman, 1991).

In the first two rounds of data collection, evaluators paid a stipend to respondents for participating in the evaluation. Under the revised data collection, we will continue this practice, using the same stipend amounts as in the past. Evaluators will offer a monetary incentive of \$20 to teachers and principals of both CSR program and non-CSR program schools. For the schools participating in the field based component, each school will receive an honorarium of \$200 to be used for purposes such as the purchase of books for the school library. The use of incentives as part of the overall strategy to increase response rates is discussed further in section B3.

A10. Assurance of Confidentiality

WestEd/COSMOS staff will comply with the Privacy Act for all individual and institutional data collected in the LACIO. The evaluators will carefully handle all data in a responsible manner so they are not seen by or released to anyone not working on the project. The evaluation team will ensure all data are reported in a summary fashion so no specific individual or school may be identified. Finally, the evaluation team will maintain all data in secure and protected files that do not include personally identifying data.

The evaluators will not collect any information that would identify individual participants. Therefore, the evaluation team will not reference participants by name or position title. WestEd will communicate an explicit statement regarding confidentiality to any and all participants. Similarly, the student achievement data extracted from the ED databases are aggregate school-level data and do not contain records of individual students.

A11. Justification of Sensitive Questions

WestEd will not ask questions that are of a sensitive nature.

A12. Estimates of Hour Burden

The estimated annual response burden is 10,774 person-hours. This total represents the sum of the estimated burden for all portions of the evaluation. Exhibit 3 aggregates the estimated total hours and costs to participants of this study. WestEd estimated the hourly rates of pay from the California Department of Education Financial Data.

Exhibit 3

Aggregate Annual Respondents and Hour Burden

Task	Number of Respondents	Hour Burden	Monetary Burden
Sampling/Gaining Cooperation	1,500	1,500	\$62,000
School Survey	13,600	8,976	\$272,646
Field-Based Study	240	240	\$9,200
State and District Survey	115	58	\$3,016
TOTAL	15,455	10,774	\$346,862

Sampling and Gaining Cooperation

Thirty sites in the field-based component have already participated in the evaluation. They have indicated their expectation that the study will continue. For the newly participating high schools, at the outset, WestEd will initiate the data collection process with phone calls to principals and district administrators to recruit schools. WestEd estimates that explaining the nature of the study and securing permission to collect data could take up to an hour on average to complete.

The number of principals shown in Exhibit 4 corresponds to the number of schools WestEd will recruit in the study (500 CSR program schools and 500 non-CSR program schools). Because WestEd intends to select the treatment and comparison schools in pairs from the same district, the number of districts is estimated at 500. This number could vary slightly for the following reasons. First, in many small districts, a comparison school will not be available, so WestEd will choose a comparison school from a neighboring district, slightly increasing the number of participating districts. Second, it is possible that in the random selection of CSR schools, WestEd could select more than one school per district – especially in large districts – potentially reducing

the number of participating districts. Due to the potential variability in this number both up and down depending upon the characteristics of school selection, 500 is a good estimate for the number of districts participating.

Exhibit 4

Estimated Annual Burden for Sampling and Gaining Cooperation

Task	Type of Respondent	Number	Time Estimate (Hour)	Total Hours	Hourly Rate	Estimated Cost of Burden
Sampling Tasks	District Administrators	500	1	500	\$52	\$26,000
Gaining Cooperation	School Principals	1000	1	1000	\$36	\$36,000
TOTAL		1,500	-	1,500	-	\$62,000

School Surveys

For the nationwide survey of schools, WestEd will request both CSR program and non-CSR program schools to participate. In each school, the principal and up to 25 teachers will complete a 40-minute survey questionnaire. Not all teachers will have such a large staff, so we have estimated that, on average, 15 teachers per school will complete the survey. With an estimated response rate of 85 percent, the potential number of total respondents will be 13,600 with a total of 8,976 person-hours. The breakdown of burden for teachers and principals for the school surveys is detailed in Exhibit 5.

Exhibit 5

Estimated Annual Burden for Respondents to School Surveys

Respondent	Total Sample Size	Estimated Response Rate	Number of Respondents	Time Estimate (Hour)	Total Hours	Hourly Rate	Estimated Cost of Burden
Principal	1,000	85%	850	.66	561	\$36	\$20,196
Teacher	15,000	85%	12,750	.66	8,415	\$30	\$252,450
TOTAL	16,000	85%	13,600	-	8,976	-	\$272,646

Field-Based Study

The targeted sample of schools selected for the field-based study consists of 20 CSR program and 20 non-CSR program schools. Interviews of the principal and three teachers at each school site are expected to last one hour each. Other data collection activities at the school site

such as classroom observations are not expected to add to the burden of the participants. The evaluators will conduct in-person interviews with two district administrators at the corresponding 20 districts for each CSR/comparison school pair. These interviews will last approximately one hour each. The evaluation team will also interview two officials from each state corresponding to each CSR/comparison school pair for approximately one hour.

Exhibit 6

Estimated Annual Burden for Participants in Field-Based Study

Respondent	Number of Respondents	Time Estimate (Hour)	Total Hours	Hourly Rate	Estimated Cost of Burden
Principal	40	1	40	\$36	\$1,440
Teacher	120	1	120	\$30	\$3,600
District Administrator	40	1	40	\$52	\$2,080
State Official	40	1	40	\$52	\$2,080
TOTAL	240	-	240	-	\$9,200

On-line Survey

In the past, LACIO used telephone surveys to gather information from all 50 state and 65 district officials. Telephone interviews were needed in order to develop the appropriate rapport with respondents and probe for varying interpretations of policy. However, the evaluators established this rapport during past data collection, so the data gathering method will change to a more efficient one, an on-line survey. Such a change reduces the burden on state and district officials. The totals for the on-line survey appear in Exhibit 7.

Exhibit 7***Estimated Annual Burden for Participants in On-line Survey***

Respondent	Number of Respondents	Time Estimate (Hour)	Total Hours	Hourly Rate	Estimated Cost of Burden
District Administrator	65	.5	33	\$52	\$1,716
State Official	50	.5	25	\$52	\$1,300
TOTAL	115	-	58	-	\$3,016

A13. Estimate of Cost Burden to Respondents

Respondents will range from assistant superintendents to teachers. The hourly rate for each respondent is outlined in section A12. There are no other additional respondent costs aside from those outlined in section A12.

A14. Estimate of Annual Cost to the Federal Government

The total cost for the evaluation is \$8,387,479 over five years. The average yearly cost is \$1,677,496. Most of the costs for the evaluation are incurred in years 2, 3, and 4 as data collection efforts commence.

A15. Program Changes or Adjustment

This request is for the revision of an existing collection. Specifically, the burden is increasing slightly, and two new research questions have been added to address the need to gather in-depth information about Title I secondary schools that are engaged in reform efforts after CSR program funding ended. A program change of 8,207 is due to the addition of two new questions on the survey forms and increased number of respondents.

A16. Plans for Tabulation and Publication of Results

The study will produce annual reports that will illuminate the progress that CSR schools made on state assessments in comparison to similar schools; the extent to which CSR reforms are effective in diverse settings; the role district and state policies played in facilitating CSR; the nature of well-implemented and sustained comprehensive reform; and the issues confronting high school reform. The timeline for the publication of these findings is outlined in Exhibit 8. The analyses for each data collection method are detailed below.

Exhibit 8

Schedule for Dissemination of Study Results

Activity/Deliverable	Due Date
<u>Final Report</u>	
Draft outline	September 28, 2007
Revised outline and summary of initial findings	October 31, 2007
First draft of report	November 30, 2007
Second draft of report	January 31, 2008
Third draft of report	March 31, 2008
Final report	June 30, 2008

Student Achievement Data

WestEd will obtain school-level student achievement data for CSR program and non-CSR program schools selected for this study from the ED databases recently compiled for ED. These databases contain achievement data from nearly every school in the country.

Because states use different tests, comparisons of student achievement are problematic. The variability in norms from different test publishers makes the comparison of absolute performance across states difficult. Using ED databases should ensure consistent data. However, if there are inconsistencies across states, WestEd will conduct overall achievement level analyses using meta-analytic techniques (Raudenbush, Fotiu, & Cheong, 1999). If there are inconsistencies over time in how data are reported, WestEd will compute standardized scores for the analysis.

School Survey Analysis

The survey will focus on research questions related to the effectiveness of CSR at the school level. Analysis will provide descriptive statistics of the extent to which each of the CSR components is implemented and provide information on the progress of schools over time. The inclusion of comparison schools from the same states and districts will enable analyses of the contribution of CSR to student outcomes.

The first step in the analysis will include computing extensive univariate statistics, plotting distributions, performing data transformations where indicated by skewed distributions, and identifying extreme outliers. Univariate screening is not sufficient to identify all errors in the data. Thus, WestEd will use cross-tabulation techniques for discrete variables measured on a nominal or ordinal scale, or for continuous variables grouped to make them discrete, to determine if relationships exist that are known not to be true or logically inconsistent. For continuous variables, a scattergram program will determine linear or non-linear relationships between variables and identify outliers. After the data are cleaned, WestEd will use exploratory

and confirmatory techniques to establish the validity and reliability of the multi-item measures. Because many of the items are measured on a nominal or ordinal scale, Muthén's (1984) approach to exploratory and confirmatory factor analysis with dichotomous and/or ordinal indicators is appropriate.

Before analyzing data across the sample, WestEd will analyze data from each sampled CSR and comparison school to determine the consistency among respondents, using cross-tabulation and correlational techniques. Differences in perceptions about school-level activities, whether associated with CSR or not, may relate to outcomes.

WestEd will address research questions that speak to the influence of state, district and school variables using multi-level modeling techniques (Bryk & Raudenbush, 1992; Goldstein, 1987). Using student achievement as the outcome variable, WestEd will estimate an unconditional, three-level, hierarchical model to decompose the total variation into three components –variation across states, variation across districts within states, and variations between CSR and comparison schools. Further, WestEd will calculate intraclass correlation coefficients to summarize the proportion of total variation that can be attributed to states, districts and schools. These intraclass correlation coefficients will be informative from a policy standpoint.

WestEd will use these procedures to simultaneously estimate school-level, district-level, and state-level models to examine how characteristics measured at higher organizational levels (*i.e.*, state) influence characteristics at lower organizational levels (*i.e.*, district, school).

Field-Based Component Analysis

The planned field-based study analysis will include both within- and across-case issues. Cases are defined as the CSR program school, a non-CSR program school, the district, and the state. Researchers will conduct one visit for each case under the reused data collection. WestEd will prepare reports in a common format to ensure all relevant elements of the conceptual framework are captured.

The site reports will provide data related to the relationship of the CSR intervention, its level of implementation, and policies of district and state, including those beyond the direct focus on CSR outcomes. The field-based studies will integrate information from the documents reviewed, interviews, focus groups, and observations held on site. As such, they can serve as stand-alone documents as well as the basis for the cross-site analysis.

To conduct the cross-case analysis, the evaluation team will use a “replication” logic, with the 20 cases (40 schools) being considered analogous to 20 different experiments in comprehensive school reform. The replication logic is the same as that used in cross-experiment

analysis, driven by logical argument and counter-argument, using evidence assembled from the individual cases. The overarching replication logic would predict that all 20 cases showed similar patterns reflecting the same generic framework of the CSR program, *e.g.* the conceptual framework. Support for such direct replications should explain deviations from the framework because of implementation failures, and not flaws in the theoretical logic of flow of events (Bickman, 1987). Multiple deviations that cannot be attributed to implementation conditions are grounds for rejecting the original framework, and developing a modified one. There might be “sub-” replication patterns, for example among schools implementing the same CSR model, where one might expect to observe the same micro-patterns in implementation and outcomes.

The inclusion of 20 comparison schools in this quasi-experimental design will result in increased ability to explain the effect of context on reform processes, implementation and outcomes. Collection of data in non-CSR sites explicitly addresses “sub-” replication patterns related to contextual conditions, as reflected in state and district policies and practices. Substantively, when CSR to non-CSR comparisons are made, a critical threat to any interpretation of results is that the non-CSR schools may not necessarily be equated to “non-treatment” sites because of contextual factors such as state and district policies. For example, the non-CSR schools may implement practices or curricula that reflect the same properties as some of the 11 components. Analysis of CSR versus non-CSR cases will focus on finding differences between sites that explain differences in student achievement trends.

Combination of Survey and Field-Based Component Analyses

Survey and field-based component data will first be analyzed separately. A qualitative cross-case analysis will compare the plausibility of different explanations for achieving comprehensive school reform, using replication logic to seek corroboratory patterns in the data. Similarly, a quantitative analysis of the survey data will reveal potential patterns associated with CSR implementation and outcomes represented by specific quantitative models. Researchers can group data by categories of interest and locate relevant data effectively. The results of both analyses can then be compared, with the comparisons driven by salient themes from the research questions and emerging themes suggested by the results. The degree to which the two independent analyses coincide or differ will form the basis for the main conclusions for the entire proposed evaluation.

Second, to the extent that the two analyses differ, and to the extent that the evaluation team believes the field-based study analysis has produced alternative explanations worthy of further analysis, WestEd will use the field-based study results to develop additional quantitative models

to be tested with the survey data. The results of the additional tests will then form the basis for additional conclusions for the entire proposed evaluation.

A17. Approval to Not Display OMB Expiration Date

WestEd is not requesting an exemption from displaying the expiration date.

A18. Explanation of Exceptions

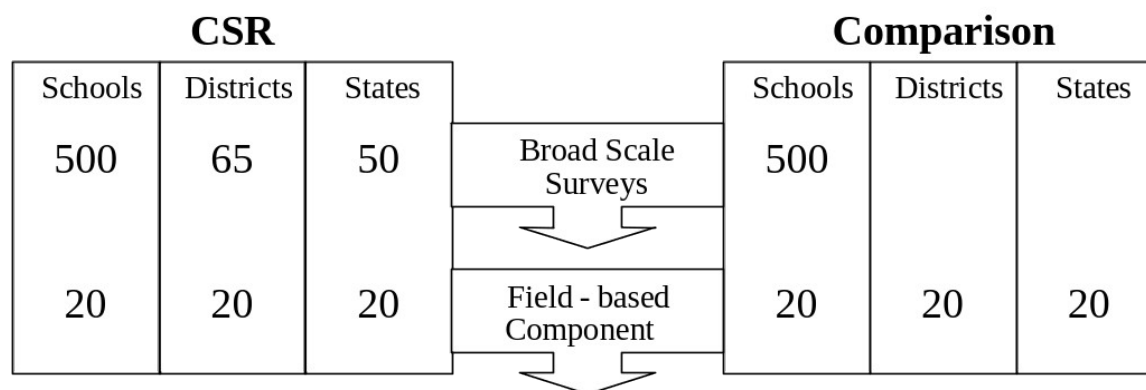
This collection of information involves no exceptions to the Certification for Paperwork Reduction Act Submissions.

PART B. DESCRIPTION OF STATISTICAL METHODS

B1. Respondent Universe

Exhibit 9 illustrates the relationship among the school samples for the mail and online surveys, and the field-based study.

Exhibit 9



Component Samples

As mentioned above, the 500 CSR program schools selected to participate in the *Quantitative Analysis of Student Achievement* and the *Survey of Reform Activities at CSR Program and non-CSR Program Schools* represent approximately one fifth of the approximately 2,000 schools that received CSR program awards in FY2002. The number of administrators, teachers, and district and state officials included in the sample are outlined in section A12 above.

Power analysis

The sample size will have sufficient power to detect differences with small to medium effect sizes. To determine the appropriate sample size, WestEd conducted a power analysis. According to Cohen, when an investigator anticipates a certain effect size, sets a significance criterion, and then specifies the amount of power desired, the investigator can determine the necessary sample size (Cohen, 1988). These four parameters of statistical inference—power, significance criterion, sample size, and effect size—are so interrelated that when three of them are fixed, the fourth is determined and can be calculated.

Using a common statistical program, the evaluation team calculated the necessary sample size, given that we expect a small effect size (.20), with a significance criterion of .05, and with varying power level of .80 (the agreed-upon minimum power level), .85 and .90. This calculation assumed a directional (one-tailed) alternative to the null hypothesis, in that we expect the CSR schools to exhibit higher assessment scores.

As the power level increases, the sample size increases. According to Exhibit 10, a sample size of 310 schools in the CSR program and non-CSR program groups is necessary to ensure the minimal power level. A sample size of 500 schools for each group was decided upon by the evaluation team to provide a more optimal power level between .85 and .90.

Exhibit 10

Sample size and statistical power

effect size	.20 (“small”)	.20 (“small”)	.20 (“small”)
alpha	.05	.05	.05
power	.80 (minimum)	.85	.90
necessary sample size	620 (310 each)	722	858

Selecting comparison schools

WestEd will select appropriately matched non-CSR comparison schools for the larger-scale survey and the field-based study using a regression-based approach to match schools according to predicted student performance. Within the district of each CSR school selected to participate in the study, WestEd will construct a School Equivalency Index (SEI) using data from ED databases. The index will be a composite measure of a school's background characteristics weighted by the portion of student performance that can be explained by those characteristics. WestEd will ask schools scoring most closely to CSR schools on the SEI to participate as comparison schools.

B2. Procedures for Data Collection

The required instrumentation for LACIO derives from a conceptual framework indicating how, in principle, Comprehensive School Reform should work at any given school. As such, the framework points to the hypotheses tested by LACIO and to the relevant variables in LACIO’s data collection.

Program theory models are a common way of expressing such frameworks. A program theory model usually depicts a presumed causal flow of events—potentially covering inputs,

outputs, and outcomes—linking sequences of actions that eventually lead to relevant outcomes. For most federal programs, the program theory models usually assume that the program-supported activity serves as an intervention (the “actions”), and the program goals serve as the desired endpoints (the “outcomes”) to be attained.

Need for a Distinctive Program Theory Model

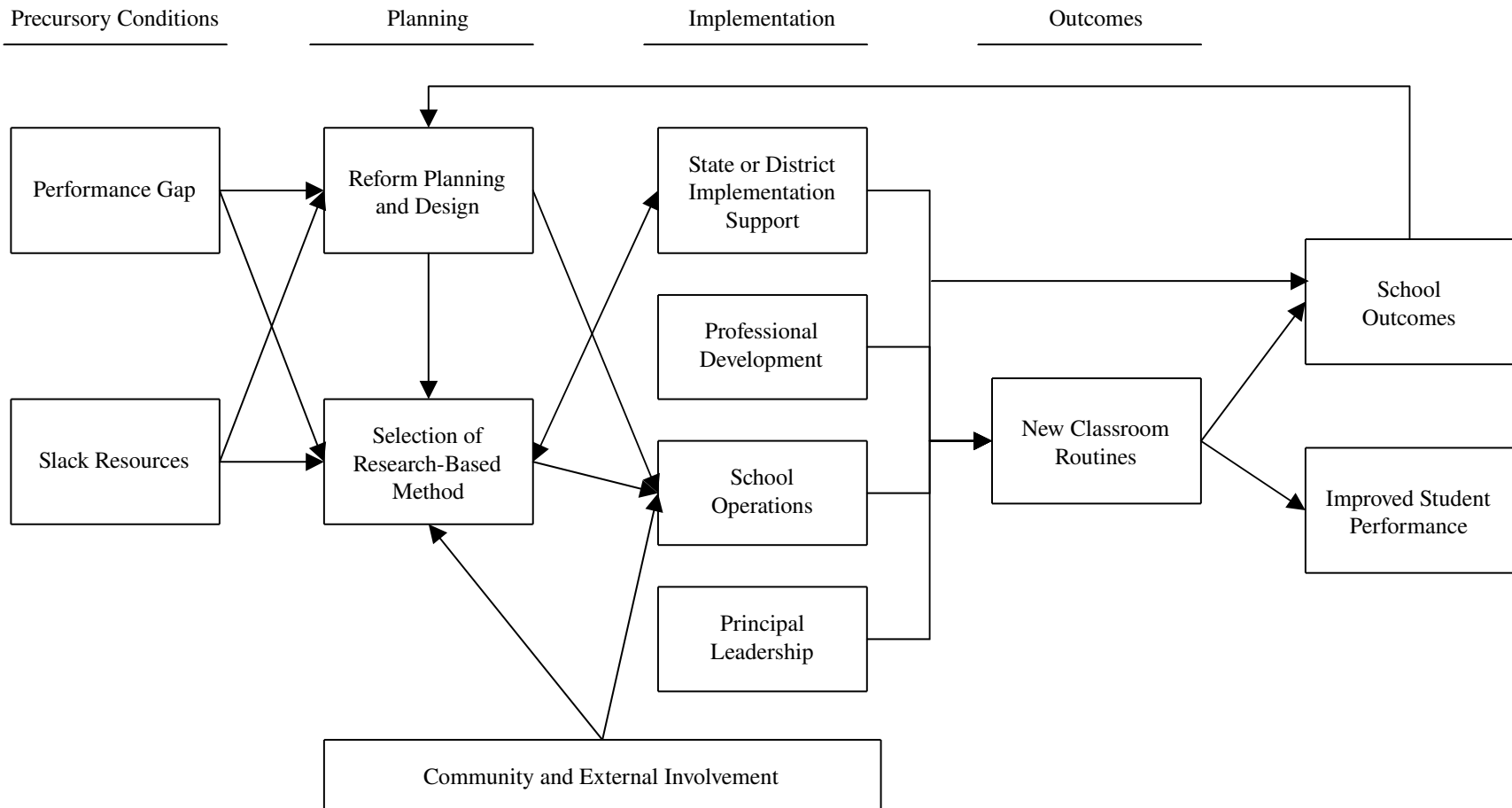
At the same time, by definition, comprehensive school reform goes well beyond the single, one-time-only intervention represented by most conventional program theory models. Successful reform is in fact an ongoing organizational process, possibly never-ending. Its goals embrace *continual* school improvement aimed at *continually* improving student performance. Such dynamic processes have not generally been represented by the traditional “input-output” program theory model.

Further, although the pre-eminent outcome of interest is improved student performance, other reform outcomes also seem conceptually relevant. A successfully reformed or reforming school seems to produce desired organizational outcomes, such as the attainment of: 1) what might be called a “performance-oriented” school; 2) whole-school alignment of values, resources, and internal organization; and 3) the empowerment of school staff, the community, and even students, to lead reform efforts on a continuing basis.

The program theory model shown in Exhibit 11 attempts to capture these two facets of comprehensive school reform - its dynamic nature and its student as well as organizational outcomes. The program theory model reflects the conceptual framework that will guide LACIO's data collection and analysis plans and is described further below.

Exhibit 11

Program Theory Model of Comprehensive School Reform Implementation and Outcomes Guiding This Evaluation



The CSR Program Theory Model

Representation of Comprehensive School Reform as a Dynamic Process. Viewing Exhibit 11 from left to right, the program theory model is organized according to a general linear sequence of events: *precursory conditions*, *planning*, *implementation*, and *outcomes*. The outcomes include the two desired types: *organizational outcomes* (new classroom routines and school outcomes) and changed *student performance*.⁵

However, closer examination reveals that the program theory model is not totally linear. The model includes a deliberately circular relationship, in which new classroom routines, school outcomes and improved student performance are not considered endpoints. Rather, their emergence represents the results of a school's ongoing reform efforts and also serves as the stimulus for continuing or modifying these efforts in the future. In turn, the new reform efforts produce a newer cycle of outcomes, again also serving as a stimulus for more reform efforts. In this manner, the program theory model captures the dynamic nature of reform.

The 11 CSR Program Components Legislated by P.L. 107-110. In addition, one of the major parts of the model captures the activities and features that mark CSR, according to P.L. 107-110 (*No Child Left Behind Act of 2001*). Some of these 11 components are hypothesized to have started during the *planning* phase and then to have extended into the *implementation* phase of a school's reform process. They include: the reform planning and design, which includes the identification of goals and benchmarks, and the planning for annual evaluations (CSR components 2, 4, and 9); and the selection of proven methods or strategies, representing effective practices having support from scientifically based research, along with sufficient staff support (CSR components 1, 11, and 5); and the community and external involvement. The remaining CSR components are listed as implementation variables.

Other Important Conditions Affecting Reform Outcomes. Despite their diversity, the 11 CSR program components do not cover at least five other important conditions that, according to previous research, affect reform outcomes. By including these five conditions, the program theory model is therefore intended to represent a more complete rendition of a hypothesized reform process. As a result, LACIO will assess schools for their fidelity to the CSR program components as well as their fidelity to these other conditions, described next.⁶

⁵ Note that stating these outcomes does not imply that they can only occur in a positive direction. LACIO may reveal negative as well as positive outcomes, and in this sense the logic model should not be considered biased in the positive direction.

⁶ This use of the term "fidelity," and the plan to collect data testing schools' fidelity both to CSR program components and to other conditions readily supported by previous research, comes directly from discussions between LACIO team and the U.S. Department of Education at a briefing on March 29, 2002.

First, Stringfield and Datnow (1998) found that, when schools implement research-based methods, the most important predictor of subsequent success is the extent to which a school has properly and thoughtfully selected the method to match its school improvement needs. Conducting a needs assessment and reviewing various candidate methods might both be part of such a process.

Second, schools are not independent entities but are part of district and state systems. Especially with the advent of test-based accountability systems, schools are increasingly being influenced by district and state policies. Besides accountability systems, other directly influential state or district policies can include such actions as the setting of: standards for what students should know and be able to do, by grade and academic subject; matriculation and graduation requirements, including the use of end-of-course examinations; strategic or school improvement planning requirements to incorporate district or state practices and objectives; and teacher certification and professional development requirements. The program theory model captures the potential district and state influence in several ways, ranging from the initial motivating conditions for initiating reform in the first place to the role in planning and designing reform to support during implementation.

Third, the bulk of educational research has repeatedly pointed to the importance of the principal as an instructional leader in producing school improvement and improving student performance. The principal strongly affects reform through such actions as: establishing the academic direction of the school; hiring teachers and defining professional development priorities; allocating resources in a manner consistent with the needs of reform; and engaging parents and the community in the reform process. The program theory model therefore directs attention to the role of the principal and the actions taken by the principal to support reform.

Fourth, because reform is to be schoolwide, a recent evaluation of an earlier cohort of CSR schools (COSMOS, forthcoming) has strongly suggested the need for rearranging some aspect of whole-school operations as an explicit signal of the reform. Without such a signal, staff, parents, students, or the community may assume that reform is limited to certain grades, subjects, or classrooms, effectively diluting the true meaning of reform. The needed type of signal could involve changes in the schedule of the school day (e.g., block scheduling or changed starting and ending hours for the school day); the assignment of teachers to additional roles such as mentoring or coaching; the formation of new leadership teams (e.g., schoolwide advisory committee); or any of a number of other organizational changes – as long as the changes are visible and affect virtually the entire school.

Fifth, the CSR components do not specify any particular change in the classroom, but curricular, instructional, or classroom management changes must occur if student performance is to improve. The particular classroom changes are likely to be linked to the nature of the proven

methods or strategies implemented by the school, but other management changes, including changes in class size, also may occur. Whatever the change(s), a further requirement is that all classrooms be affected, a process that can occur incrementally over time.

Precursory Conditions. In addition to all of the CSR and non-CSR conditions leading to reform outcomes, the reform also starts somewhere. The program theory model therefore contains precursory conditions hypothesized to lead to reform: a combination of a perceived or real *performance gap* combined with the availability of some *slack resources*, typically taking the form of external federal, state, or private funds, such as CSR program funds. Existing research repeatedly established the importance of these two aspects of the program theory model (Timperly & Robinson, 2001). Therefore, the main hypothesis for the LACIO evaluation is that schools would not have embraced reform processes in the absence of *both* of these precursory conditions.

Behaviorally anchored survey items

The surveys used in the evaluation represent the desire to measure respondents' actual behaviors rather than attitudes or beliefs. WestEd will maximize the survey data by asking about a school's, district's, or state's actual behaviors and conditions. Such items may be compared to the traditional design of survey items, emphasizing subjective ratings and responses from individual respondents that are vulnerable to soliciting "socially desirable" responses.

When activity or behavioral items are used (e.g., "*Which academic subjects are covered by goals or benchmarks for student achievement?*"), in principle the items are more amenable to external corroboration, and respondents are more likely to give their most accurate response. Although the entire survey procedure still consists of self-reported data, the use of such behavioral items should increase the quality of the survey data.

All instruments used in the evaluation use items that measure behaviors and are evidence-based, rather than attitudes and expectations. The research questions central to the evaluation drive the selection of survey items. Because the LACIO research questions focus on identifiable actions of schools and their impact as the result of CSR program implementation, the survey administered to participants in this evaluation will focus strictly on behaviors. In essence, the survey becomes a survey of behaviors that result from the implementation of CSR rather than attitudes about its implementation. As Fowler (1993) notes, there are right or wrong answers to questions posed around such objective states, allowing researchers to accurately describe people and/or events. This type of survey will be not only more informative but also less subject to response bias. Questions that require responses of factual information are less sensitive to the tendency of respondents to answer attitudinal questions in a socially desirable manner.

B3. Methods to Maximize Response Rate

WestEd will use mail surveys to collect teacher and principal data from the 1,000 schools in the study and Web-based surveys to collect data from state and district officials. Though mail surveys are low cost, easy to implement, widely used, and contain no interview bias, their response rates are often low. To overcome this problem, WestEd will implement a variety of well-documented strategies shown to increase the rate of mail survey responses.

Such strategies produced high response rates for the LACIO in the first two years of data collection. The analyses presented in this *Second-Year Report* (in press) are based on a response rate of 90 percent from the 478 CSR schools that agreed to participate and 82 percent from the 481 non-CSR schools that agreed to be part of the study. An initial analysis of responses from the most recent data collection found that 88 percent of CSR schools and 78 percent of non-CSR schools provided survey data for the forthcoming *Third-Year Report*. Based on such past participation and the success of follow-up efforts, we expect an 85 percent response rate for this revised data collection.

Research on how to increase survey response is a well studied area of survey methodology (Boser & Clark, 1993; Church, 1993; Cole, Palmer, & Schwanz, 1997; Dillman, 1991, 1996; Fox, Crask, & Kim, 1988; Kanuk & Berensen, 1975), producing a variety of successful strategies. In order to keep survey response high WestEd will employ four of these strategies: (1) survey sponsorship, (2) advance contact, (3) incentives, and (4) follow-up.

The sponsorship of the survey increases the response rate (Heberlein and Baumgartner, 1978; Boser & Clark, 1996). Surveys sponsored by universities and government agencies usually have higher response rates than surveys sponsored by private companies. For government-sponsored surveys, participants feel their input in a survey can result in public policy change that will affect them.

Advance contact to participants also increases the response rate (Boser & Clark, 1996). This advanced contact can be in the form of a telephone call, postcard, or letter mailed to the participant prior to receiving the survey. Some schools may require district approval before participating in the study. In these cases, WestEd will contact district personnel and request school participation. Whenever possible, WestEd will ask districts to officially encourage the schools to participate in the evaluation.

Incentives are items offered to participants as motivation to complete a survey. They can be enclosed with the survey or promised to the participant at a later time. Enclosed incentives have a higher response rate than promised incentives (Church, 1993). Incentives can be monetary or non-monetary (movie tickets, coupons, a chance to participate in a drawing). With the exception

of follow-ups, past research shows monetary incentives are the most successful technique for improving mail survey response rates (Mangione, 1995).

Generally, larger monetary incentives yield a larger response rate (Boser & Clark, 1996). But two studies, one involving the National Survey of Family Growth (NSFG) and the other the National Adult Literacy Survey (NALS) suggest that \$20 dollars is the peak amount for an effective monetary incentive (Berlin, Mohadjer, Waksberg, Lolstad, Kirsch, Rock, and Yamamoto, 1992; Mosher, Pratt, and Duffer, 1994). Both studies found that incentives below \$20 decreased response rates, but incentives above \$20, such as \$34 and \$40 did not significantly raise response rates.

WestEd will offer the same incentives as in the original study: a monetary incentive of \$20 to teachers and principals of both CSR program and non-CSR program schools. For the 40 schools participating in the field-based component, WestEd will offer an honorarium of \$200 per school, to use for such purposes as the purchase of books for the school library. Also, as a non-monetary incentive, WestEd will offer results of LACIO findings to the schools participating in the study. Providing study results to schools is a recommended way to help offset the burden of staff completing surveys (Cole, et al., 1997).

Follow-up strategies take the form of postcards, letters, phone reminders and replacement surveys sent to participants who have not responded to the survey. Follow-up mailings that include additional surveys generate higher returns than just reminder letters (Boser & Clark, 1996).

Four weeks after the initial mailing of the survey, Duerr Evaluation Resources (DER) will telephone principals (and perhaps teachers) to follow-up with non-respondents. Wherever necessary, DER will complete the survey over the phone with non-respondents.

B4. Expert Review and Field Testing

All data collection instruments will go through a series of reviews before use in the field. The evaluation team will work closely with ED throughout the design and revision process to ensure that the instruments and items adhere to the issues raised in the legislation. The evaluation team will also share instruments and procedures with members of the TWG, who will bring their expertise as researchers and practitioners to bear on the design of the items, the burden on respondents, and the implications for data analysis.

In addition to the TWG review, the evaluation team will pilot test all data collection instruments. During these tests, which will be administered to no more than nine respondents, the team will assess item comprehension, the effectiveness of the proposed strategies for gaining cooperation, and the length of time for respondents to answer questions on the instruments. Such

information is critical for determining the burden associated with each instrument, which must be presented to all respondents before the administration of any federally sponsored research questionnaire to more than nine respondents.

B5. Statistical Consultants

Dr. Naida Tushnet at WestEd and Dr. Robert K. Yin at COSMOS Corporation are co-directing the LACIO. They bring years of experience conducting evaluations and investigating school reform activities.

Dr. Tushnet first directed a national study of exemplary Title 1 and bilingual education projects in 1976. While at the National Institute of Education, she was responsible for designing and monitoring major studies of dissemination and knowledge use. She also directed a number of national studies, including *Documentation and Evaluation of the Educational Partnerships Program* and the *Evaluation of the Star Schools Program* (both for the Office of Educational Research and Development); evaluations of the National Science Foundation's Instructional Materials Development Program and NSF's Mathematics Across the Curriculum reform of undergraduate mathematics teaching; and the evaluation of the development of Different Ways of Knowing for the Middle Grades.

Dr. Yin has directed numerous large-scale national studies, including the *CSRD Field-Focused Study*, part of the current national evaluation of the CSRD program. For NSF, he is leading the evaluation of the Urban Systemic Program, NSF's major investment in urban schools districts' math and science education, and previously directed the national evaluation of the Experimental Program to Stimulate Competitive Research (EPSCoR), covering a 15-year period in which 19 states received an average of \$3 million per year per state. Dr. Yin also led the *National Evaluation of the Local Law Enforcement Block Grant Program* and the *Evaluation of the CDC-Supported Technical Assistance Network* for community planning for HIV/AIDS prevention.

Staff assigned to the project also have experience with national studies. Jerry Bailey serves as chief statistician on virtually all WestEd (and previously, SWRL) survey studies. John Flaherty has coordinated the LACIO since 2002 and contributed to national studies of charter schools. Brooke Connolly and Jin-Ok Kim provide analytic expertise to LACIO and will continue this work under the revised data collection effort. From COSMOS, Dr. Janeula Burt coordinated data collection for LACIO and other national evaluations. She and Dr. Darnella Davis, who serves as a senior analyst for COSMOS' *Mathematics and Science Partnerships Program Evaluation* (MSP-PE), will coordinate the data collection and analysis at COSMOS for the reused data collection.

Division of labor

WestEd will be responsible for coordinating the evaluation activities and conducting the analysis of student achievement. WestEd and COSMOS will share responsibility for refining the research design and developing the data collection instruments.

DER, a small business firm, will work closely with the WestEd and COSMOS team to develop survey instruments. DER is a professional evaluation firm based in Chico, California. The firm's primary business activity since 1981 has been conducting evaluation and survey research. The firm's major responsibility is conducting the school survey including mailing the surveys, conducting follow-up contacts, and entering data. DER will also collaborate with WestEd in analyzing the survey data.

WestEd will have a key responsibility for sampling and survey data analysis, assisted by DER in the latter task. The evaluation team will divide the field-based research evenly between WestEd and COSMOS. At this point, the team anticipates that the COSMOS group will have responsibility for sites east of the Mississippi River and WestEd will conduct case study visits in sites west of the Mississippi. However, the team leaders will make site assignments after site selection, seeking the most cost effective assignments.

REFERENCES

- Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences* Second Edition
- COSMOS Corporation (forthcoming). Field-Focused Study of CSR. Bethesda, MD: Author.
- Timperly, H. S., & Robinson, V. M. J. (2001). Achieving school improvement through challenging and changing teachers' schema. *Journal of educational Change* 2 (4), 281-300
- Berlin, M., Mohadjer, J., Waksberg, J., Lolstad, A, Kirsch, I., Rock, D., and Yamamoto, K. (1992). An experiment in monetary incentives. Proceedings of the Survey Research Methods Section of the American Statistical Association, 393-398.
- Bickman, L. (1987). The functions of program theory. In L. Bickman (Ed.), Using program theory in evaluation. New Directions for Program Evaluation, No 33. San Francisco: Jossey Bass.
- Boser, J.A., and Clark, S.B. (1993). Response Rates in Mail Surveys: A Review of the Reviews. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction Service No. 356378).
- Boser, J.A., Clark, S.B. (1996). Reviewing the research on mail survey response rates: Descriptive study. Presented at the Annual Meeting of the American Educational Research Association, New York, NY.
- Bryk, A., and Raudenbush, S. (1992). *Hierarchical linear models: Applications and data analysis method*. Thousand Oaks: Sage.
- Church, A.H. (1993). "Estimating the effect of incentives on mail survey response rates: a meta-analysis." Public Opinion Quarterly, 57: 62-79.
- Cole, C., Palmer, R., Schwanz, D. (1997). Improving the mail return rates of SASS surveys: A review of the Literature. National Center for Education Statistics (ED), Washington, DC.
- Dillman, D.A. (1991). The Design and administration of mail survey. Annual Review of Sociology, 17: 225-249.
- Fowler, F. (1993). *Survey research methods*. Thousand Oaks: Sage.
- Fox, R.J., M.R. Crask and J. Kim (1988). Mail survey response rates: A meta-analysis of selecting techniques for inducing response. Public Opinion Quarterly, 52: 467-491.
- Herberlein, T.A. and R. Baumgartner (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. American Sociological Review, 43: 447-462.

Kanuk, L. and C. Berensen (1975). Mail surveys and response rates: A literature review. Journal of Marketing Research, 12: 440-453.

Mangione, T.W. (1995). Mail Surveys: Improving the Quality. Applied Social Research Methods Series, Vol. 40. Thousand Oaks, CA: Sage Publications.

Mosher, W.D., Pratt, W.F., and Duffer, A.P. (1994). CAPI, event histories, and incentives in the NSFG cycle 5 pretest. Proceedings of the Survey Research Methods Section of the American Statistical Association, 1, 59-63.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115-132.

Raudenbush, S.W., Fotiu, R.P., Cheong, Y.F. (Winter 1999). Synthesizing results from the trial state assessment. *Journal of Educational and Behavioral Statistics*, 24 (4), 413-438.

Stringfield, S., & Datnow, A. (1998). Scaling up school restructuring designs in urban schools. *Education and Urban Society*, 30 (3), 269-276.

U.S. Department of Education. (2004). *Longitudinal assessment of Comprehensive School Reform (CSR) Program implementation and outcomes: First-Year report*. Washington, DC: Author.

U.S. Department of Education. (in-press). *Longitudinal assessment of Comprehensive School Reform (CSR) Program implementation and outcomes: Second-Year report*. Washington, DC: Author.

Appendix A

Linking Research Questions and Instrument Items

Research Question	Principal Survey	Teacher Survey	State Survey	District Survey	Site Visit Protocol	Classroom Observations
Extent of Grade-Level and Content Planning and Coordinated Curriculum and Instruction						
Type of planning activities	2, 4	2, 3		10	2.2-2.3, 7.4	
Training or guidance for planning	4, 5	2, 4	2	2	2.1, 4.2	
Data used in planning	3, 5	4			7.2, 7.5	
State and district role in planning	2-5	4, 5, 6	1	1	B2	
Coordinated Curriculum and Instruction			3-4	3-4		C
Extent of Data Driven Decision Making (DDDM)						
Type of data (formative and/or summative) used	6-8, 10, 11	7-9, 12	4-7	4,6-8	2.2, 7.5	
Who uses which data	9, 10	10		5		
Training for DDDM	10, 12	11	6,12	6	7.2	
Changes as a result of DDDM	12, 35	13		8		
Role of School Improvement Plans (SIPs)						
Development of SIP	13-16	14-17	9	11	B1-B2, 7.3	
Data used in SIP	12, 15	16, 17	9	11	7.2	
Evaluation of SIP	16, 17	17, 18	8,10	9,12	10	
Dissemination of SIP		17, 18				
Changes in program as a result of SIP	35	37	8	9	10	
Extent of Professional Development (PD)						
Types of PD offered	18, 19	19, 20, 22	11	13	B4, 4	E1
Source of PD	4		2,17	14	B4	
Alignment of PD with state content standards, state curriculum, local curriculum, and/or SIPs	18	6	3-4,12	15	B2	E2-E3
Adoption of Scientific Method						
Types of reform methods implemented	23, 26, 28	25, 28	13	16	1.1-1.3, 2.1	D
Alignment of methods and state and district standards	5, 30	32			9	E1
Research basis for the reforms chosen	27, 30	32	13	16	2.1, 6.2	
Comprehensiveness of the reforms chosen	28	27, 30			1.1-1.3	
Evaluation of the reforms chosen					10	
Capacity for Reform & Sustainability						
Support for reform at the school	31, 39	33, 41		20	B3, B5, 7.1	D, E1, E4
Assistance provided to facilitate reform	32, 39	34, 41	17	21	8.1-8.2	
Barriers or limits to teacher participation in reform	33, 36, 37, 39	35, 36, 38, 39, 41	16	19,20	B1-2, 9.2, 11.2	
Roles of state and district actions in the school's implementation of reform	16, 27, 29, 31	5, 29, 31	1-2,17	1,21	B2, 9, 11.1-11.2	