



Implementation: Measuring and Explaining the Fidelity of CSR Implementation

Anja Kurki, Daniel K. Aladjem and Kevin R. Carter



AMERICAN
INSTITUTES
FOR RESEARCH®

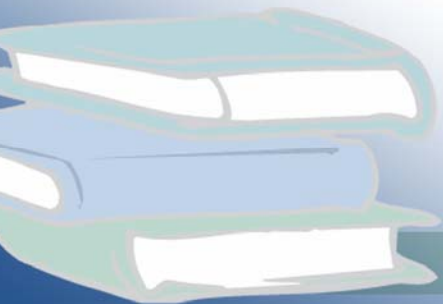
1000 Thomas Jefferson Street, NW | Washington, D.C. 20007-3835 | 202.403.5000 | www.air.org

Implementation: Measuring and Explaining the Fidelity of CSR Implementation

A paper prepared for the annual meeting of the
American Educational Research Association, Montreal, Canada, April 11–15, 2005

Anja Kurki, American Institutes for Research®
Dan Aladjem, American Institutes for Research®
Kevin R. Carter, American Institutes for Research®

The contents of this document were developed under a grant from the United States Department of Education (PR/Award Number R306S000012). The contents, however, do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.



Contents

Abstract	3
Introduction	4
Measuring Implementation in the National Longitudinal Evaluation of Comprehensive School Reform Study.....	6
How Does the Level of Implementation Vary Across Schools?	10
Do CSR Schools Have Higher Levels of Implementation Than Comparison Schools?	11
Does the Level of Implementation Vary Among CSR Model Keys?.....	14
Variation Among School Districts	15
Is Fidelity of Implementation Related to Years of Implementation?	15
Are CSR Models Implemented Comprehensively?.....	16
What Predicts the Level of and Change in CSR Model Implementation?	19
In What Level Is Variance in Implementation Located?	20
Teacher–School Models.....	21
Results	24
Year 1 of the Study (2002).....	24
Year 3 of the Study (2004).....	32
What Predicts Change in Implementation Between 2002 and 2004?.....	38
Results.....	39
Discussion	46
References	49
Appendix A: Implementation Indices	51
Calculating Implementation: Euclidean Distance.....	52
Appendix B: Propensity Scores	54
Appendix C: Survey Questions Used to Create Implementation Indices	56
Appendix D: Box Plots Showing Variation and Implementation by CSR Key and by District, Year 1	68
Implementation by CSR Key.....	68
Implementation by District.....	74
Appendix E: Significant Differences Among CSR Model Keys	80



List of Figures

Figure 1. Implementing Change.....	5
------------------------------------	---

List of Tables

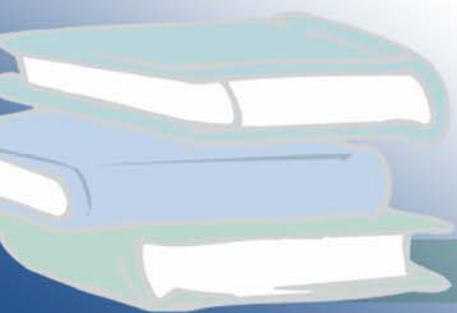
Table 1. Descriptions of Implementation Indices	8
Table 2. Number of CSR Schools and Comparison Schools Matched to Specific CSR Model.....	9
Table 3. Descriptive Statistics for Implementation Indices in CSR and Comparison Schools.....	10
Table 4. Descriptive Analysis Results: Variation in the Level of Implementation.....	12
Table 5. Do CSR Schools Have Higher Levels of Implementation Than Control Schools?.....	13
Table 6. Total Number of Indices Implemented at Average Level or Above-Average Level	17
Table 7. Number of Classroom Instruction-Related Indices Implemented at Average or Above-Average Level.....	17
Table 8. Level of Comprehensiveness in CSR and Control Schools, Year 1	18
Table 9. Level of Comprehensiveness in CSR and Control Schools, Year 3	18
Table 10. Location of Variance (Teacher vs. School vs. School District).....	21
Table 11. Regression Results for Level of Implementation, Year 2002.....	28
Table 12. Regression Results for Level of Implementation, Year 2004.....	34
Table 13. Regression Results for Change in Implementation From 2002 to 2004	42
Table B1. Number of CSR Schools and Control Schools Matched to Specific CSR Programs	55
Table C1. Survey Questions Used to Create Implementation Indices.....	56
Table E1. Significant Differences Among CSR Model Keys, Year 1	80
Table E2. Significant Differences Among CSR Model Keys, Year 3	83



Implementation: Measuring and Explaining the Fidelity of CSR Implementation

Abstract

Comprehensive school reform (CSR) is only as effective as its implementation. By using data collected for the National Longitudinal Evaluation of Comprehensive School Reform, this study explores how CSR model implementation varies as well as what factors predict CSR model implementation. We found little difference in the level of implementation between CSR model schools and their matched comparison schools, but vast differences were found in the level of implementation between (a) different components of CSR models and (b) different CSR model-implementation keys (the normative guidelines received from CSR model developers on which we based our implementation measures). The level of implementation is predicted by multiple factors, including the CSR model-implementation keys, the principal's instructional leadership, the measure of teachers' professional community, and the usefulness of the CSR developers' assistance. The change in implementation is associated with positive change in principal's instructional leadership.



Implementation: Measuring and Explaining the Fidelity of CSR Implementation

Introduction

Any comprehensive school reform (CSR) effort is only as good as its implementation. CSR models are assumed to profoundly change practices and activities in multiple dimensions, from school governance and instruction to parent and community involvement. The changes facilitated by CSR models are the source of improved student outcomes, not the mere presence of a CSR model. The likelihood of successful reform—successful implementation of a CSR model and improvement in student achievement—involves coordinated actions of multiple actors on multiple levels: students, teachers, parents, principals, and district and state administrators. Accordingly, some schools are more likely to be successful in their reform efforts than others.

The success of implementing a CSR model has been connected to multiple factors. These factors include the characteristics of the CSR model, the context (school/community/school district/state/federal) in which the reform is taking place, the process of implementing a CSR model, and the incentives for implementation. Models that are “locally grown,” or less specific or prescribed concerning curriculum and instruction, tend to be less faithfully implemented (Desimone, 2002). Similarly, an environment with conflicting or competing reforms or policies does not foster implementation with fidelity to the model, whereas a coordinated reform effort between different levels of actors (teachers, schools, school district, state) positively affects implementation of school reform (Hatch, 1998). The level of implementation has also been related to school characteristics, existing stereotypes, and norms. Schools in high-income areas with smaller proportions of minority students have been more successful in implementing school reform programs (Berends, Bodilly, & Kirby, 2002; Yonezawa & Datnow, 1999).

The process of choosing and implementing a specific CSR model is also important. A program that fits the needs of a school and is supported or chosen by principal and teachers together is likely to be better implemented (Slavin, 1999). Teachers who have supportive professional networks and common planning time to tackle problems related to school reform programs are likely to implement programs with higher fidelity (Cooper, Slavin, & Madden, 1998; Stringfield & Datnow, 1998). Although teachers are the ultimate implementers of the reform programs, the principal’s leadership skills can make a significant difference. Principals who secure adequate resources, are involved in everyday instructional decisions, and/or are charismatic leaders have been associated with higher levels of implementation (Berends, 2000; Berends et al., 2002; Berends, Kirby, Naftel, & McKelvey, 2001).

In this paper, we attempt to identify and describe some factors that promote or hinder implementation of CSR models. To accomplish this, we address four specific research questions:

1. Are schools that implement CSR models engaged in different sets of practices and activities regarding components of CSR models than schools that are not implementing CSR models?
2. Does the level of implementation vary among CSR models?
3. Does the level of implementation vary by component and by how comprehensively CSR models are implemented?

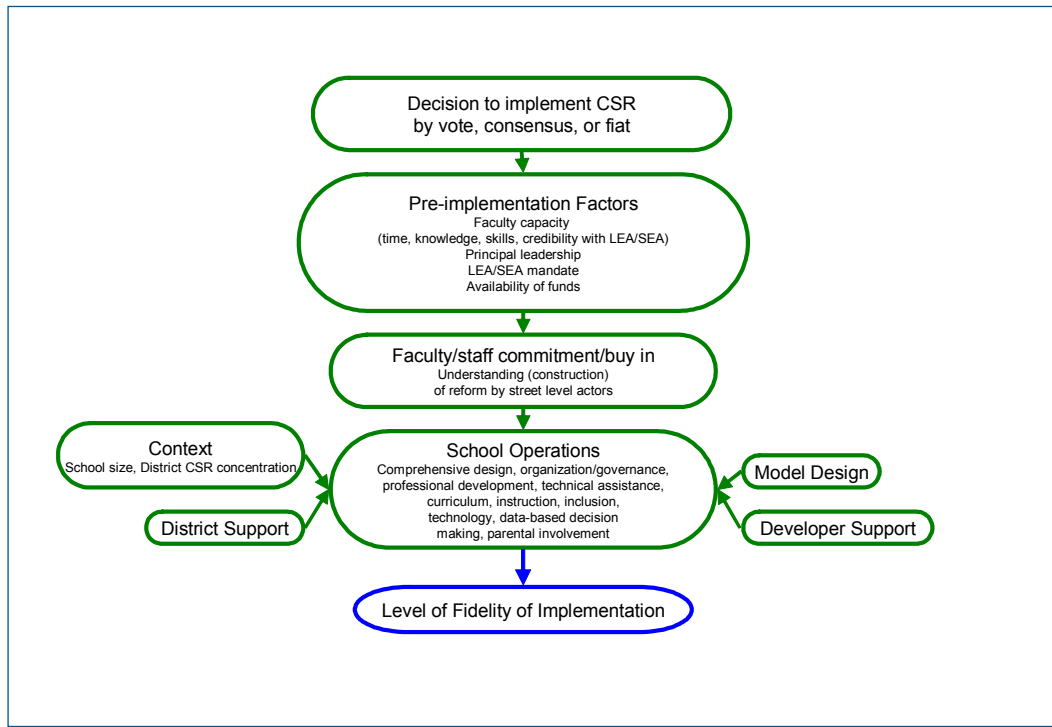


4. What factors, other than having a specific CSR model, predict the level and change in implementation?

These research questions are answered by applying a quantitative measure of CSR model implementation based on CSR model developers' survey answers, descriptive statistical analysis, and multilevel regression approach.

In answering these questions, our analysis is guided by a conceptual model shown in Figure 1.

Figure 1. Implementing Change



We assume that the fidelity of implementation is affected by CSR model design; that is, by the set of practices and activities in which a CSR model school should be engaged if the school is fully implementing a CSR program. We call this set of practices and activities, identified by different CSR developers, a CSR model key. In addition, it is hypothesized that fidelity of implementation is influenced by the CSR developer's support; support from the school district; faculty commitment and buy-in; and both teacher- and school-level characteristics, including contextual measures for schools (poverty, school size, etc.), teachers (English teacher vs. mathematics teacher), and faculty capacity (measures of professional community and experience).

Our conceptual model clearly identifies the importance of district support in relation to the level of implementation of CSR models. However, because of the small number of districts included in the study ($n = 21$) and the lack of balanced distribution of CSR programs among districts, our preliminary results use only two-level (teacher and school levels) hierarchical linear modeling (HLM) models.



Measuring Implementation in the National Longitudinal Evaluation of Comprehensive School Reform Study

Implementation of CSR has been widely studied, and the level of implementation has been measured both qualitatively and quantitatively. Qualitative measures of implementation have often been based on a combination of model developer's benchmarks (Bifulco, Duncombe, & Yinger, 2005; Bodilly, Keltner, Purnell, Reichardt, & Schuyler, 1998; Smith et al., 1998); classroom observations (Bloom, Rock, Ham, Melton, & O'Brien, 2001; Datnow, Borman, Stringfield, Rachula, & Castellano, 2003; Engelmann & Engelmann, 2004; Smith et al., 1997; Smith et al., 1998); and interviews of teachers, principals, and staff. Quantitative measures of implementation have relied on measures based on principals' and teachers' self-reported survey answers and on indirect measures of implementation, such as the number of years a school has been implementing a CSR model. (Barnes, Camburn, Kim, & Rowan, 2004; Berends et al., 2001; Berends et al., 2002; Cook et al., 1999; Kirby, Berends, Naftel, & Sloan, 2001; Supovitz & May, 2003).

Like most studies of CSR model implementation, the National Longitudinal Evaluation of Comprehensive School Reform (NLECSR) study applies both qualitative and quantitative measures of implementation. The *qualitative analysis* of implementation captures the process and comprehensiveness of implementation from the points of view of principals, teachers, and staff in a subsample of study schools. Our approach for the *quantitative analysis* is quite different. We measure implementation as *fidelity*: the extent to which the CSR model of interest is delivered to the intended recipients in the intended way (Aladjem, 2003). The approach that we have developed (summarized in Appendix A) is based on the idea that to measure the fidelity of implementation, we should measure what schools are doing and compare that with what CSR model developers consider to be "full" implementation. The challenge of measuring implementation is finding the difference between the positive, empirical reality of school life and the normative vision of CSR model developers.

We operationalized this process by asking CSR model developers to fill out the same survey instruments (principal and teacher surveys) as our survey respondents, as if they were a fully implementing school. We compared the survey answers from our principals and teachers to the answers of their respective CSR model developer. Then we calculated the distance between the ideal developer-specified implementation (the developers' answers) and the actual implementation taking place in schools (the principals' and teachers' answers). Squared Euclidean distance was used to calculate the difference between developers' and principals'/teachers' answers, and the distance measure was transformed to percentage of implementation to increase intuitive interpretation of the results (see Appendix A for a full description of the methods used to calculate implementation indices). Thus, our measure of implementation can be understood to measure how fully a school is engaged in those practices that a school should be engaged in during full implementation of a specific CSR model.

Calculating the distance between the developers' ideal answers and principals'/teachers' actual answers produced implementation scores for schools that are implementing CSR models. But we also needed implementation measures for our comparison schools that are not implementing CSR models. We wanted to compare schools implementing a specific CSR model to schools that we know are not implementing any CSR model. This comparison lets us see whether schools that have adopted CSR models are indeed engaged in a different set of practices. This comparison is one valid method, in addition to longitudinal analysis that includes data for schools before and after they adopted a CSR model, to test whether CSR models do change how schools operate. We do not have data for the vast majority of schools before they



adopted a CSR model; therefore, the longitudinal analysis is impossible, and it is crucial that comparison schools are included in the analysis.

Consequently, we had to determine what specific CSR model the comparison schools would be likely to implement, if they were to choose to implement one. We decided to use a propensity score approach to match comparison schools to a specific CSR model. We calculated propensities of being a CSR school for all schools included in the study. Then, based on their propensities of being CSR schools, we matched comparison schools and CSR schools within each school district. Thus, if a comparison school had a similar propensity of being a CSR school as a school implementing CSR model “C,” the comparison school would be “assigned” to CSR model C to calculate its level of implementation (see Appendix B for a full description of this method). As a result, we are able to compare the level of CSR model implementation between the schools in the study that implemented CSR and comparison schools that have similar characteristics.

- ◆ To create our implementation indices, we asked developers of CSR models to fill out both our principal and teacher surveys. Seven CSR developers returned the surveys. Therefore, only schools implementing one of those seven CSR models and respective comparison schools could be included in the implementation analysis.¹

The principal survey was administered three times (once for each year of the study); the teacher survey was administered twice, during years 1 (2002) and 3 (2004) of the study. In this paper, we focus on the full set of possible implementation indices, using the information provided by the CSR model developers in the teacher surveys. Thus, the paper concentrates on implementation measures for years 1 and 3 of the study, and the possible increase/decrease in implementation between years 1 and 3 of the study. The implementation indices used in descriptive analysis are school-level measures, based on information from principal or teacher surveys (or both): the implementation indices calculated for each teacher who returned a survey have been aggregated up to school level to create a school-level measure of implementation. In the two-level HLM model including teacher and school levels, we use teacher-level implementation indices.

The indices that were created, based on the principal and teacher survey instruments, are closely aligned with the 11 components of the federal CSR grant program (20 USC 6516). Our implementation indices include seven general categories and 13 specific implementation indices (see Table 1).

¹ In addition, because a small number of school districts have ardently adopted CSR models, we have lost most comparison schools in these districts, and consequently lost most of the CSR schools as well.

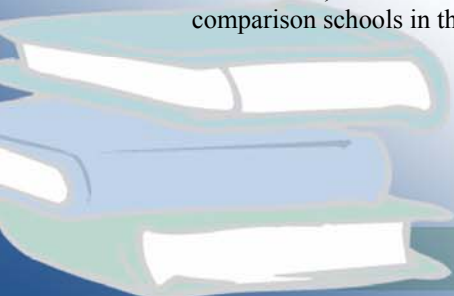


Table 1. Descriptions of Implementation Indices

<p>Governance</p> <p>Shared Decision Making: To what extent decision-making authority in a school is shared among faculty, staff, and administrators according to the CSR program developer’s ideal of shared decision making. This index uses both principal and teacher survey items.</p> <p>School Organization: To what extent a school’s organization matches the CSR program developer’s ideal of school organization. This index uses principal survey items.</p>
<p>Technology</p> <p>Use of Technology in Classrooms: To what extent the ways teachers use technology in their classroom match the CSR program developer’s ideal.</p>
<p>Parent/Community Involvement</p> <p>To what extent a school’s actions regarding parent–school communication and community involvement match the CSR program developer’s ideal. This index uses both teacher and principal survey items.</p>
<p>Professional Development</p> <p>Emphasis of Professional Development (PD): To what extent the PD received by teachers matches the CSR program developer’s ideal regarding the type (all grade-level teachers collectively vs. noncollectively) and emphasis of the PD.</p> <p>Engagement in Informal PD: To what extent teachers’ engagement in informal PD matches the CSR program developer’s ideal.</p>
<p>Assessment</p> <p>Influence of Assessments: To what extent different types of assessments influence students’ grades, grouping decisions, adjusting curriculum, etc., according to the CSR program developer’s ideal.</p> <p>Use of Assessments: To what extent the teacher is using classroom assessments according to the CSR developer’s ideal.</p>
<p>Organization of Teaching/Classrooms</p> <p>Inclusion: To what extent non-native English speakers and students with disabilities are mainstreamed in general education classes according to the CSR program developer’s ideal. This index uses both teacher and principal survey items.</p> <p>Student Grouping: To what extent students are taught in similar or mixed ability groups and how often these groups change according to the CSR program developer’s ideal.</p> <p>Time Scheduled for Teaching: To what extent the frequency and length of instructional time matches the CSR program developer’s ideal.</p>
<p>Instruction</p> <p>Curriculum: To what extent do teachers teach mathematics or English topics according to the CSR program developer’s ideal.</p> <p>Pedagogy: To what extent do teachers engage in different instructional activities according to the CSR program developer’s ideal.</p>

Note. See Appendix C for a full list of survey items used in creation of these indices.



This paper uses these indices to show how the levels of implementation vary (a) between CSR and comparison schools, (b) among schools having implementation scores calculated based on surveys returned by developers of a specific CSR (CSR model key), (c) by time (phase of implementation), and (d) over time. In addition, these indices are outcomes in two-level HLM analysis where the level of implementation in schools will be predicted by teacher- and school-level variables. Tables 2 and 3 provide basic descriptive statistics for all implementation indices included in the analysis.

Table 2. Number of CSR Schools and Comparison Schools Matched to Specific CSR Model

CSR model	Year 1		Year 3	
	CSR schools	Comparison schools	CSR schools	Comparison schools
CSR Model A	33	52	29	48
CSR Model B	13	19	12	16
CSR Model C	27	33	22	28
CSR Model F	85	93	80	87
CSR Model D	9	18	8	15
CSR Model E	12	11	9	10
CSR Model G	6	4	6	2
Total	185	230	166	206

Table 2 clearly shows that the sample is imbalanced. CSR Model F is overrepresented, whereas CSR Model D, CSR Model E, and CSR Model G are underrepresented. Surprisingly, the proportion of schools using different CSR models in the sample closely mirrors the proportion of schools using these CSR models in the United States.

Table 3 provides basic descriptive information about the level of implementation in different components of CSR models, as captured by the implementation indices included in the study. Table 3 shows that the level of implementation is similar in CSR schools and comparison schools, but the level of implementation varies significantly by index. For instance, the level of implementation is generally high regarding practices related to classroom instruction, such as Curriculum, Time Scheduled for Instruction, Pedagogy, and Use of Assessments. Similarly, the size of standard deviations (*SDs*) for implementation indices varies: the less highly implemented indices, such as Use of Technology, Inclusion, and Parent/Community Involvement, also tend to have higher standard deviations, thus implying that practices included in these indices are not implemented according developers' guidelines.



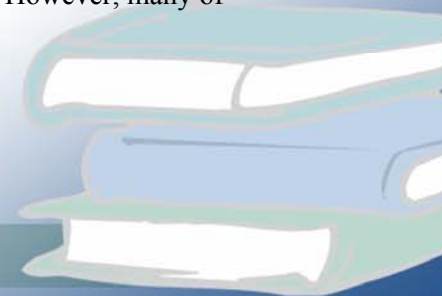
Table 3. Descriptive Statistics for Implementation Indices in CSR and Comparison Schools

Index	Year 1 CSR schools					Year 1 Comparison schools					Year 3 CSR schools					Year 3 Comparison schools				
	Mean	SD	n	Min	Max	Mean	SD	n	Min	Max	Mean	SD	n	Min	Max	Mean	SD	n	Min	Max
Implementation of School Organization	0.72	0.167	164	0.15	1.0	0.70	0.180	192	0.12	1.0	0.56	0.186	130	0	1.0	0.59	0.195	176	0	1.0
Shared Decision Making	0.75	0.090	164	0.29	0.90	0.70	0.094	192	0.37	0.91	0.74	0.114	141	0.10	0.98	0.75	0.086	175	0.46	0.96
Inclusion	0.66	0.148	164	0.16	1.0	0.65	0.166	191	0.00	0.91	0.77	0.127	139	0.16	1.0	0.78	0.110	176	0.24	1.0
Parent/Community Involvement	0.69	0.119	164	0.40	0.95	0.69	0.116	192	0.41	1.0	0.72	0.126	141	0.38	0.96	0.71	0.134	176	0.38	0.96
Professional Development: Emphasis and Type	0.72	0.084	163	0.47	0.91	0.73	0.085	192	0.41	0.91	0.74	0.137	132	0	0.91	0.73	0.116	164	0.23	1.0
Use of Technology in Classrooms	0.54	0.144	155	0	0.88	0.53	0.139	182	0.18	0.85	0.58	0.159	128	0.19	0.96	0.54	0.168	156	0	0.89
Student Grouping	0.87	0.088	163	0.53	1.0	0.87	0.077	192	0.60	0.99	0.91	0.085	132	0.5	1.0	0.89	0.095	163	0.45	1.0
Curriculum	0.82	0.074	140	0.52	0.99	0.80	0.078	153	0.54	0.97	0.82	0.078	97	0.51	0.97	0.80	0.088	128	0.44	1.0
Time Scheduled for Instruction	0.86	0.126	163	0.36	0.99	0.84	0.143	191	0.39	0.98	0.87	0.153	132	0.29	1.0	0.85	0.168	165	0.24	1.0
Engagement in Informal Professional Development	0.72	0.100	155	0.35	0.96	0.72	0.103	177	0.43	0.96	0.72	0.114	131	0.25	0.93	0.74	0.138	154	0.34	0.95
Influence of Assessments	0.76	0.054	155	0.61	0.91	0.76	0.070	176	0.56	0.95	0.77	0.069	126	0.52	0.93	0.74	0.096	153	0.45	0.94
Use of Assessments	0.87	0.043	163	0.67	0.95	0.87	0.044	192	0.74	0.98	0.87	0.061	132	0.59	1.0	0.86	0.067	164	0.62	1.0
Pedagogy	0.88	0.096	151	0.62	0.99	0.87	0.110	172	0.58	0.99	0.89	0.116	109	0.46	0.99	0.87	0.134	146	0.46	0.99

Some CSR developers did not provide survey answers required to create all the indices included in this paper. For example, CSR Model F developer did not return the separate survey for English teachers; therefore, Pedagogy and Curriculum indices could not be developed for English teachers. CSR Model D developer did not answer survey questions needed to calculate indices for Inclusion, Parent/Community Involvement, Curriculum, Professional Development, Influence of Assessment, and Pedagogy. In addition, we do not have survey answers regarding Use of Technology, Inclusion, and Parent/Community Involvement for CSR Model E, and CSR Model B did not provide needed information regarding Curriculum. Consequently, analyses of these indices are incomplete, with lower numbers of total cases.

How Does the Level of Implementation Vary Across Schools?

We have conceptualized implementation of CSR models for descriptive analysis in three different ways. First, schools that are implementing CSR models should engage in activities encouraged/required by a specific CSR model more often than comparison schools that are not implementing the CSR model. For instance, it is assumed that on average, CSR Model F schools would more often regroup their students for reading language/arts sessions than schools that are not implementing CSR Model F. However, many of



the practices promoted by CSR models—such as changes in school governance, PD for teachers, and including non-English-speaking students and students with disabilities in regular classrooms—have been mainstreamed and are currently in place in a majority of schools. Therefore, we suggest a second measure of implementation, the standard deviation of our implementation indices. The concept of using this measure is simple: even if a group of CSR model schools is not implementing CSR in a significantly higher level (higher mean), they should do it more consistently (lower standard deviation). In other words, it is possible that some comparison schools have adopted most of the practices supported by CSR models without implementing a CSR program, while most comparison schools have adopted only a few of these practices. As a result, mean values of implementation indices for schools implementing CSR models and comparison schools might be similar, but the standard deviation would be higher for comparison schools.

Our third measure of implementation captures the idea of the “comprehensiveness” of CSR models. Each CSR model is supposed to address all 11 components of the federal CSR grant program (20 USC 6516). One of the 11 components defines comprehensive design for CSR models:

Comprehensive design—A comprehensive design for effective school functioning integrates instruction, assessment, classroom management, professional development, parental involvement, and school management. By addressing needs identified through a school needs assessment, it aligns the school’s curriculum, technology, and professional development into a plan for schoolwide change. The ultimate goal of this design is to enable all students to meet challenging State content and student academic achievement standards.

Accordingly, we created a new categorical variable to measure comprehensiveness based on the 13 implementation indices created for the study. This variable has a range from 0 (no index is implemented at average or above-average level) to 13 (all indices are implemented at average or above-average level).

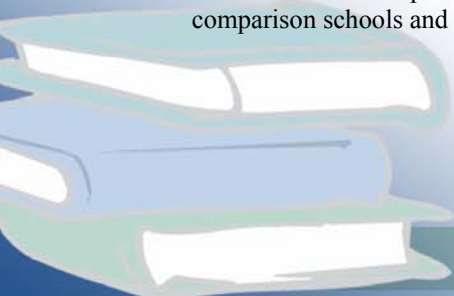
Do CSR Schools Have Higher Levels of Implementation Than Comparison Schools?

We first conducted simple descriptive analyses to see how implementation varies (a) between CSR schools and their matched comparison schools, (b) among different CSR models, and (c) among different school districts (see column 1 in Table 4 and Appendix D for box plots). Because we have only a few indices for CSR Model D, and CSR Model G includes only a few schools, these schools were dropped from the analysis in testing for statistical significance.

According to results of two-way analysis of variance (ANOVA),² conducted by using data from years 2002 and 2004 (year 1 and year 3 of data collection), the mean level of CSR model implementation is similar in CSR and comparison schools for almost all indices. The exceptions in 2002 are Parent/Community Involvement and Inclusion indices. As a group, CSR schools implement Parent/Community Involvement with lower fidelity than comparison schools, whereas Inclusion is implemented with higher fidelity in CSR schools.

The results for the year 2004 were similar; only a few significant differences existed.

² The simple two-way ANOVA model included variables for CSR keys (identifying both CSR and comparison schools that have implementation indices based on the same developer) and variables identifying CSR and comparison schools and their interaction.



CSR schools were more often engaged in Shared Decision Making than comparison schools. In addition, CSR schools had significantly higher fidelity of implementation regarding Use of Technology in Classrooms. Finding only a few significant differences between CSR schools and comparison schools is understandable: five CSR models are included in the implementation analysis, and the activities and practices recommended by the models are different. Therefore, differences between comparison schools and combined measures of CSR models are likely to be diluted. Consequently, it is important to see whether differences in implementation between schools implementing a specific CSR model and their matched comparison schools exist.

Table 4. Descriptive Analysis Results: Variation in the Level of Implementation

Implementation index	Difference between all CSR schools and matched control schools		Difference among CSR keys		Difference in implementation index over time	Variation in level of implementation by phase of implementation		Difference in implementation among school districts: CSR schools only	
	Year 1	Year 3	Year 1	Year 3		Year 1	Year 3	Year 1	Year 3
Governance									
Shared Decision Making		(+)*	***	***				***	***
School Organization			***	***	Decrease***			***	**
Technology		(+)**	***		Increase***			***	*
Parent/Community Involvement	(-)*	(-)*	***	***	Increase***		**	***	***
Professional Development (PD)									
Emphasis of PD					Increase*			***	
Participation in PD			***	***	Increase**				***
Assessment									
Influence of Assessments			***			**		***	
Use of Assessments			***	***			***	***	
Organization of Teaching/Classrooms									
Inclusion	(+)*		***	***	Increase***			***	**
Student Grouping			***	***	Increase***		**		**
Time Scheduled for Teaching			***	***		**	***	***	***
Instruction									
Curriculum			***	***				***	*
Pedagogy			***	***	Increase***			***	***

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$; (+)/(-) = CSR schools have higher or lower level of implementation, respectively.



To test whether schools implementing specific CSR models have higher levels of implementation than comparison schools, we compared CSR schools to their matched comparison schools (using two-way ANOVA, including CSR model keys and actual CSR school designation). Because only a few schools are implementing some of the CSR models included in the analysis, we report detailed results for only the four CSR models implemented in the largest number of schools: CSR Model A, CSR Model B, CSR Model C, and CSR Model F (see Table 5).

Table 5. Do CSR Schools Have Higher Levels of Implementation Than Control Schools?

Index	CSR Model A		CSR Model B		CSR Model C		CSR Model F	
	2002	2004	2002	2004	2002	2004	2002	2004
Governance								
Shared Decision Making			Lower**					
School Organization								
Technology								
Parent/Community Involvement			Lower***					
Professional Development								
Emphasis of PD								
Engagement in Informal PD								
Assessment								
Influence of Assessments	Higher**							
Use of Assessments	Higher*							
Organization of Teaching/Classrooms								
Inclusion								
Student Grouping								
Time Scheduled for Instruction								
Instruction								
Curriculum								
Pedagogy	Higher*							

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

The results are surprising. Schools implementing CSR models did not systematically have higher means than their matched comparison schools for the implementation indices. Only CSR Model A schools had higher levels of implementation than their matched comparison schools in 2002. CSR Model A schools implemented the Influence, Use of Assessments, and Pedagogy with higher fidelity than their matched comparison schools. These differences did not continue to exist in 2004. CSR Model B schools implemented Shared Decision Making and Parent/Community Involvement with lower fidelity than their comparison schools. Indeed, no significant differences existed between CSR schools and their matched comparison schools in 2004.



Additional descriptive analysis (descriptive statistics for CSR model schools and their matched comparison schools) did not reveal any obvious trends regarding standard deviations of the implementation indices. That is, the standard deviations were not consistently larger or smaller for comparison schools matched with a specific CSR model. This implies that schools implementing a specific CSR model, even if having a mean similar to their comparison schools, are not implementing models with higher consistency.

Several plausible explanations exist for these results. First, it is possible that most schools, whether or not they implement a CSR model, are engaged in similar activities. That is, many practices initiated and promoted by CSR models may have become mainstreamed. If so, then the relevant question is how these practices (which are included in the implementation indices) become implemented, i.e., which developer's key is successful in positively affecting student achievement. Second, the ANOVA analysis does not control for other factors that could affect the level of implementation. Therefore, HLM analysis may reveal additional differences between CSR and comparison schools. Finally, it is possible that our sample of CSR-implementing schools for some of the programs are biased toward known "low implementers." Developers of CSR Model B, for instance, recognized some schools included in our sample as low implementers, and CSR Model B schools are the only ones having significantly lower levels of implementation than their matched comparison schools.

Does the Level of Implementation Vary Among CSR Model Keys?

The above analysis compared schools implementing specific CSR models to their matched comparison schools. However, it is also important to know whether different CSR models have varying levels of implementation. The focus of this analysis is not to compare the implementation levels of schools implementing different CSR models but to understand how developers' keys (survey answers given by developers) affect the level of implementation. It is possible that some developers' implementation keys more closely mirror the practices already existing in schools and thus are "easier" to implement. To examine how the implementation keys affect the level of implementation, we included both the CSR schools and their comparison schools in the analysis. That is, all schools that have implementation scores based on CSR developers' survey answers are included in the analysis. Descriptive results based on two-way ANOVA show that the level of implementation varies tremendously among CSR model keys.

Only one index, Emphasis of PD, does not have significantly different mean levels among different CSR model keys (see column 2 in Table 4 and Appendix D for box plots). Interestingly, the levels of implementation are higher for all schools (those implementing and not implementing CSR models) in most practices related to classroom instruction: Curriculum, Pedagogy, Grouping of Students, and Influence and Use of Assessments. The standard deviations for these indices are also smaller (when compared to the mean levels of indices), implying that instructional practices and activities are (a) considered to be important to be implemented with fidelity and (b) consistently implemented. The exception is the Time Scheduled for Instruction; this index shows clear differences among different CSR model keys. Indices about School Governance, Use of Technology, Inclusion, and Engagement in Informal PD have larger mean differences among the CSR model keys, and the standard deviations for these measures are higher, implying that not all schools implement these components with high fidelity.

To examine how specific CSR keys differ from each other, Tukey-adjusted means from ANOVA analysis were used³. In 2002, schools implementing CSR Model E and CSR Model F model keys were consistently among the higher implementers (Parent/Community Involvement and Use of Assessments

³ We focused on general trends. For examples of tables of specific outcomes, see Appendix E.



were the exception for CSR Model F). Schools using CSR Model A and CSR Model B model keys consistently have lower level of implementation than most other CSR models keys in Shared Decision Making, Inclusion, Parent/Community Involvement, and Use of Technology in Classrooms. The CSR Model A model key also has a lower level of Pedagogy implementation. Schools using CSR Model C model key also have lower levels of implementation concerning Parent/Community Involvement, Student Grouping, and Time Scheduled for Instruction.

The same trends continued in 2004. Schools implementing CSR Model E and CSR Model F model keys were consistently among the high-level implementers. Schools using CSR Model A model key were still lower implementers concerning Parent/Community Involvement, Engagement in Informal PD, Influence of Assessments, and Pedagogy, while schools using the CSR Model B model key were still among the lowest implementers regarding Shared Decision Making, Inclusion, and Parent/Community Involvement. Schools using the CSR Model C model key continued to implement Student Grouping, Time Scheduled for Instruction, and Parent/Community Involvement with lower fidelity to the model.

As mentioned above, the meaning of “faithful implementation” or “implementation with fidelity” varies for each CSR model included in the analysis. Therefore, it is possible that the combination of implementation keys (surveys returned by CSR model developers) and our way of calculating implementation (the distance between developer’s answers and the actual answers given by principals/teachers) makes implementation of some CSR model keys more difficult. In addition, different dimensions of CSR are more central for different CSR models; these factors are also likely to affect the values of our implementation indices. For instance, one central dimension in the CSR Model B is School Governance, especially Shared Decision Making. Accordingly, the CSR Model B model key requires a higher level of shared decision-making authority at the school level, and these practices are not completely mainstreamed. Consequently, the CSR Model B model key for Shared Decision Making requires *change* and is more difficult to implement. Schools may also decide to implement the aspects of the model selectively. If the main problem in the school is perceived to be Instruction, the Parent/Community Involvement dimension of the CSR model may receive less attention and, as a result, be implemented at a lower level.

Variation Among School Districts

To examine variation in CSR model keys, we conducted another ANOVA model. For this analysis, we included 16 school districts. Vast differences exist in the levels of CSR model key implementation among school districts (see column 5 in Table 4 and box plots in Appendix D). Indices measuring classroom activities, such as Curriculum, Pedagogy, Use and Influence of Assessments, and Student Grouping, tend to be much more consistently implemented (with high means and smaller standard deviations). The exceptions are Use of Technology and Inclusion. The latter measures how well schools implement mainstreaming of non-English-speaking students and students with disabilities. Again, measures not directly related to instruction, such as Governance and Parent/Community Involvement, are implemented at lower levels and with less consistency. These results closely mirror those relating to CSR model keys because, in many school districts in our sample, only one or two CSR models predominate. Consequently, district differences pick up differences between CSR model keys.

Is Fidelity of Implementation Related to Years of Implementation?

Two different analyses were conducted to see whether fidelity of implementation is related to the length of implementation. First, we studied whether the level of implementation varies with the years of CSR model implementation. We created a variable with the following categories: 1–2 years of implementation,



3–5 years of implementation, and more than 5 years of implementation. We tested whether the mean values of our implementation indices are significantly different between these three categories. The results were somewhat surprising: only two of the indices had significantly different means in 2002. In 2004, four of the indices had significantly different means (results based on Tukey-corrected means; see Table 4). In 2002, both Time Scheduled for Instruction and Influence of Assessments had higher mean values in schools that had implemented a CSR model for 5 or more years than in schools with 3–5 years of implementation. Moreover, in regard to Influence of Assessments, schools with 5 or more years of implementation were also better implementers than schools with 0–2 years of implementation. In 2004, schools with 5 or more years of implementation had higher means for Parent/Community Involvement, Student Grouping, and Time Scheduled for Instruction than schools with 3–5 years of implementation. However, this relationship to time of implementation was reversed for Use of Assessments.

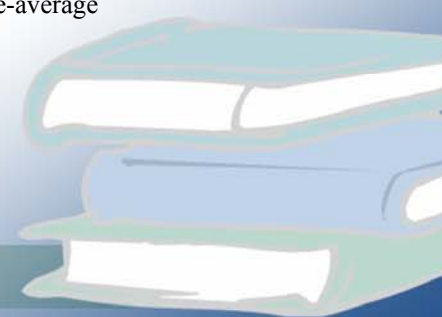
Second, we explored whether the level of implementation changed between 2002 and 2004. According to our results, the level of implementation goes up for most indices in both CSR and comparison schools, implying that practices promoted by many CSR models became increasingly popular over time. The largest positive changes took place for Inclusion (mainstreaming students) and Use of Technology. Four of the 13 indices showed either no change or a decrease. The index for School Organization decreased over time. This measure captured organizational changes that often take place in CSR schools, such as having a coach or coordinator who helps to implement CSR models. This result makes intuitive sense. Salaries of coaches or coordinators and other supports are often paid by a CSR grant that lasts 3 years; many schools may not be able to keep the coach or coordinator when the grant money is gone. For the Use of Assessments, Influence of Assessments, and Curriculum, the level of implementation was already high for the first year of the study, making additional increase more difficult (i.e., the ceiling effect). It is also possible that these three indices are largely determined by school districts; without a change in district policy, a change in school practices is less likely.

Are CSR Models Implemented Comprehensively?

A main point about CSR models is that they are supposed to be comprehensive—changing activities and practices ranging from school governance to instruction. Our analysis so far has demonstrated vast differences in the level of implementation among different dimensions of CSR models. Therefore, we decided to examine further how comprehensive the implementation of CSR models is in our sample of schools. For this analysis, comprehensiveness was defined as follows: a school that implements at average or above-average level in all 13 CSR implementation indices would have achieved fully comprehensive implementation.⁴ This comprehensiveness measure does not measure fidelity of implementation, because different implementation indices are not weighted according to their importance for a specific CSR model design. Rather, our measure is based on a broader understanding of comprehensiveness. That is, schools implementing a CSR model should not be picking and choosing some CSR dimensions to implement; instead, schools should have relatively high levels of implementation of all CSR dimensions.

Our results show a wide range of comprehensiveness in implementation. During year 1 of the study, only one school implemented all 13 indices at average or above-average level; during year 3 of the study, none

⁴ Each implementation index is changed to a dummy variable, taking a value of 1 if a school has an average or above-average level of implementation; otherwise, the dummy variable takes a value of 0. Then a count variable was created by summing up the values of each dummy variable to see how many indices each school implements at the average or above-average level. A fully implementing school would have an average or above-average implementation for all 13 indices, i.e., a value 13 for the count variable.



of the schools had attained fully comprehensive implementation (see Tables 6 and 7). These results are not surprising, because different CSR models concentrate on different dimensions in their design. Therefore, we created another measure that included only classroom-related practices, which are central elements in most if not all the CSR models included in the study. In this measure, 26 (14%) schools in year 1 and 23 (13%) schools in year 3 were comprehensively implementing all five classroom-related indices; 32 (28.6%) and 46 (40.3%) schools at year 1 and year 3, respectively, were implementing two or fewer indices.

Table 6. Total Number of Indices Implemented at Average Level or Above-Average Level

Number of indices implemented	Year 1		Year 3	
	Frequency	Percent	Frequency	Percent
0	—	—	1	0.26
1	14	3.4	4	1.0
2	17	4.1	13	3.3
3	24	5.8	31	7.9
4	41	9.9	27	6.9
5	53	12.8	45	11.5
6	60	14.5	55	14.0
7	56	13.5	60	15.3
8	60	14.5	61	15.6
9	53	12.8	37	9.4
10	28	6.8	28	7.1
11	7	1.7	23	5.9
12	2	0.5	6	1.5
13	—	—	1	0.26

Table 7. Number of Classroom Instruction-Related Indices Implemented at Average or Above-Average Level

Number of indices implemented	Year 1		Year 3	
	Frequency	Percent	Frequency	Percent
0	10	2.4	19	4.9
1	50	12.0	40	15.1
2	91	22.0	102	41.1
3	116	28.0	122	31.1
4	102	24.6	74	18.9
5	46	11.1	35	8.9



To further analyze comprehensiveness of implementation, we divided schools into three categories—low, middle, and high comprehensiveness⁵—and tested whether significant differences exist between CSR schools and comparison schools regarding comprehensiveness of CSR implementation (see Tables 8 and 9). According to the chi-square test, the distributions for CSR and comparison schools are significantly different for year 1 (at 0.1 level) and year 3 (at 0.01 level); more CSR schools than comparison schools belong to the “high” category of comprehensiveness.

Table 8. Level of Comprehensiveness in CSR and Control Schools, Year 1

Level of comprehensiveness	Year 1 (2002)	
	Control school	CSR school
Low comprehensiveness	<i>n</i> = 60 (26.1%)	<i>n</i> = 36 (19.5%)
Middle comprehensiveness	<i>n</i> = 128 (55.9%)	<i>n</i> = 101 (54.6%)
High comprehensiveness	<i>n</i> = 42 (18.3%)	<i>n</i> = 48 (26.0%)

Table 9. Level of Comprehensiveness in CSR and Control Schools, Year 3

Level of comprehensiveness	Year 3(2004)	
	Control school	CSR school
Low comprehensiveness	<i>n</i> = 64 (31.2%)	<i>n</i> = 35 (21.3%)
Middle comprehensiveness	<i>n</i> = 129 (62.9%)	<i>n</i> = 108 (65.9%)
High comprehensiveness	<i>n</i> = 12 (5.9%)	<i>n</i> = 21 (12.8%)

Our preliminary analysis shows several interesting results that will be tested in multilevel HLM analysis. First, schools implementing CSR are not engaged more often in activities recommended by CSR model developers than schools not implementing CSR models. A few exceptions exist, but in general, this is the trend in the preliminary descriptive analysis. The result of no differences between CSR schools and their matched comparison schools could be caused by the mainstreaming of practices initiated and promoted by CSR models or by the fact that CSR models include common school practices that need to be implemented qualitatively differently. Our measure of implementation cannot capture qualitative differences in practices implemented in CSR and comparison schools (although the implementation indices often capture differences in emphasis and frequency of activities and practices). For instance, both CSR and comparison schools may have high levels of Shared Decision Making, but the actual process of

⁵ Schools with low comprehensiveness are implementing fewer than four dimensions; schools with middle comprehensiveness implement four to eight dimensions; and schools with high comprehensiveness implement more than eight dimensions.



how decisions are made may be quite different. Measurement of these qualitative differences is a problem faced by all self-reported survey measures of implementation.

Second, large differences exist in the level and consistency (measured by standard deviation) both in implementation of different dimensions of CSR models and among different CSR models' keys. These differences can be explained at least partly by different foci of CSR models. That is, some programs promote more general governance change while others concentrate on changing instruction-related practices. Nevertheless, our measure of implementation is constructed to capture fidelity of implementation: how closely schools match the CSR model developer's concept of ideal implementation. Thus, the results show that instruction-related practices are more highly and consistently implemented than practices not related to instruction, such as Shared Decision Making, Use of Technology, and Parent/Community Involvement. An interesting exception is Inclusion (mainstreaming students). Although Inclusion is an instruction-related practice, it was not very highly or consistently implemented in 2002 but was increasingly implemented in 2004. In addition, the significant differences among CSR model keys, together with the fact that few significant differences exist between CSR schools and their matched comparisons, support the idea that some of the CSR model keys are more difficult to implement than others. In other words, some CSR models require practices that are not commonly done in schools; consequently, these CSR models have lower levels of implementation. If this is the case, it is important to establish whether CSR model keys have differentiated effects on outcomes such as student achievement.

Third, the level of implementation of a few practices, such as Time Scheduled for Instruction, Use and Influence of Assessments, and Student Grouping, is related to the phase of CSR implementation. Implementing a CSR program is not likely to be a linear process, in which the level of implementation steadily increases over time. Rather, different components of CSR models may be implemented at different levels at different times: school governance and organization-related changes may take place early on, whereas changing instruction is likely to take a longer time.

Finally, the analysis clearly shows that CSR is not very comprehensively implemented. In 2002, 26.0% of CSR schools and 18.3% of comparison schools had comprehensive implementation. The percentage of schools engaged in comprehensive implementation decreases over time. Only 12.8% of CSR schools and 5.9% of comparison schools included in the analysis have comprehensive implementation (high comprehensiveness) in year 2004.

What Predicts the Level of and Change in CSR Model Implementation?

Two different analytical models are used to examine what predicts the level of and change in CSR implementation. The first analytical model has two levels⁶: teachers and schools. Here, our interest is to see how much of the variance in implementation is located within and between schools and what explains variance in these levels. This analytical model will be analyzed cross-sectionally for all implementation indices, using data from years 1 and 3. The second analytical model examines what predicts change in implementation (an outcome derived by subtracting year 1 teacher-level implementation indices from year 3 teacher-level implementation indices).

⁶ We opted to have a two-level model instead of a three-level model because of data problems at the district level (few districts, and very disproportionate allocation of CSR models and control schools among districts) and due to the small amount of variance to be explained at the district level for many implementation indices.

In What Level Is Variance in Implementation Located?

To see how variance is partitioned between different levels, an unconditional three-level HLM-model was conducted (see Table 10).

In 2002, the variance at the teacher-level ranged from 31.1% (Inclusion) to 96.4% (Student Grouping). In addition to the Student Grouping index, both PD indices, both Assessment indices, as well as the Technology and Curriculum indices have more than 80% of their variance at the teacher level. Results for these indices are similar to results of Berends (2000), who found that approximately 80–90% of variation in implementation resides at the teacher level. However, one would expect less teacher-level variance for some implementation indices that measure school-level phenomena. Accordingly, Inclusion, Time Scheduled for Instruction, and Shared Decision Making indices have less variance at the teacher level. Surprisingly, the Pedagogy index that captures practices used in the classrooms by teachers had a large proportion of its variance at the school level (30%). This finding could be due to schoolwide and/or districtwide PD regarding instruction as well as prescriptive curriculums.

Most indices had 10% or more variance at the school level. The exceptions are the Assessment indices, Emphasis of PD, Student Grouping, and Curriculum. The variance at the school level for Parent/Community Involvement, Shared Decision Making, Pedagogy, Inclusion, and Time Scheduled for Instruction ranged from 18.5% to 44.4%. Only Pedagogy and Inclusion indices had more than 10% of their variance at the district level (26.0% and 27.5%, respectively).

In 2004, most of the implementation indices had 75% or more of their variance at the teacher level. As in 2002, the exceptions were Inclusion, Time Scheduled for Instruction, and Pedagogy. For Pedagogy and Time Scheduled for Instruction, the amount of variance at the teacher level decreased but increased at the school and district levels; for Inclusion, however, the variance at the teacher level increased.

Taken together, these results imply that some practices captured by implementation indices are driven by district or school policies. The practices included in the Pedagogy and Inclusion indices are most clearly school and district driven (most of the variance would be either in school or district level). Activities included in Shared Decision Making and Parent/Community Involvement vary across the schools but less across districts. Interestingly, Curriculum- and Assessment-related activities, which arguably could be determined by school districts (use of specific curriculum and programs), have most of their variance at the teacher level. In general, a trend of increased variance at the teacher level seems to exist over time.



Table 10. Location of Variance (Teacher vs. School vs. School District)

Outcome	Teacher		School		School district	
	2002	2004	2002	2004	2002	2004
Governance						
Shared Decision Making	68%***	78.8%***	23.4%***	12.6%***	8.6%**	8.6%**
Technology	82.45***	88.0%***	10.6%***	7.5%***	7.0%***	4.4%*
Parent/Community Involvement	77.5%***	81.7%	18.5%***	18.0%***	4.0%*	0.3%
Professional Development						
Emphasis of PD:	90%***	86.8%***	7.5%***	11.0%***	1.6%*	2.1%*
Engagement in Informal PD:	82.5%***	77.7%***	13.8%***	18.0%***	3.6%**	5.0%***
Assessment						
Influence of Assessments	84.8%***	85.0%***	9.9%***	13.5%***	6.7%**	1.6%
Use of Assessments	90.3%***	692.0%***	6.3%***	6.2%***	3.4%**	1.8%
Organization of Teaching/Classrooms						
Inclusion	31.3%***	49.1%***	41.2%***	33.4%***	27.5%***	17.3%**
Student Grouping	96.4%***	98.5%***	3.4%**	1.2%	0.2%	0.3%
Time Scheduled for Teaching	49.4%***	30.0%***	44.4***	57.8%***	6.2%***	11.2%***
Instruction						
Curriculum	81.3%***	88.1%**	9.8%***	6.7%	8.9%*	5.2%
Pedagogy	43.5%***	33.9%***	30.5%***	37.9%***	26.0%***	28.2%***

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Teacher–School Models

A cross-sectional model for years 2002 and 2004 was conducted to examine what predicts the level of implementation. The outcomes of these analyses are teacher-level implementation indices that mirror the school-level indices used in the descriptive analysis above (see Table 1). According to our conceptual model (see Figure 1), we can predict the level of implementation by the following variables:

1. Two sets of dummy variables: the first set for different CSR models' keys (identifying CSR schools and their matched comparison schools); the second set for schools implementing specific CSR models (school-level variables)
2. Variables controlling for school context: Student Enrollment, Challenging Environment Scale, Percentage of Students Receiving Free/Reduced-Price Lunch, Percentage of Non-English-Speaking Students, Existence of a CSR Coach/Coordinator in the Schools, School with Middle Grades, and School's Adequate Yearly Progress (AYP) status (school-level variables)



3. Principal leadership: Principal's Instructional Leadership Scale (teacher-level variable); a dummy variable recognizing new principals with 1 year or less experience in their current school (school-level variable)
4. Teacher background: a dummy variable identifying teachers with only a bachelor's degree; a dummy variable identifying new teachers with 1 year or less in their current school; dummy variable distinguishing between English and mathematics teachers; a dummy variable identifying teachers who supported adoption of the CSR model vs. those who did not support, were not given an opportunity to influence the decision, or were not at the school at the time of adoption (teacher-level variables)
5. Usefulness of Developer's Support Scale (teacher-level scale)
6. Years of CSR implementation: dummy variables for 3–5 years and for more than 5 years of implementation (school-level variables)

The teacher- and school-level model can be given as:

Teacher level

$$Y_{ij} = \beta_{0j} + \beta_1 Context_{ij} + \dots + \beta_4 Context + \beta_5 Leadership_{ij} + \beta_6 Developer + v_{ij}$$

Where:

Y_{ij} = implementation index for i th teacher in j th school

$Context_{ij}$ = teacher-level contextual variables for i th teacher in j th school

Leadership = teachers' report on principal's instructional leadership

Developer = teachers' report on usefulness of developer's assistance

v_{ij} = a random error term that is independently and identically distributed across teachers within school

β_{0j} = average level of implementation index for school j

$\beta_{1j} \dots \beta_{4j}$ = effect of teacher-level contextual variables within school j

β_{5j} = effect of principal's instructional leadership (teacher report) on implementation index within school j

β_{6j} = effect of usefulness of developer's assistance (teacher report) on implementation index within school j



School level

$$\beta_{0j} = \gamma_0 + \gamma_1 CSRkey + \dots + \gamma_5 CSRkey + \gamma_6 CSRprogram + \dots + \gamma_{11} CSRprogram + \gamma_{12} Context + \dots + \gamma_{18} Context + \gamma_{19} NewPr incipal + \gamma_{20} Phase1 + \gamma_{21} Phase2 + v_j$$

Where:

CSRkey = dummy variables for CSR model keys

CSRprogram = dummy variables for implemented CSR model

Context = school-level context variables

Leadership = principal with 1 year or less of experience in the current school

Phase1 = dummy for 3–5 years of CSR model implementation

Phase2 = dummy for more than 5 years of CSR model implementation

v_j = a random error term that is independently and identically distributed across teachers

γ_0 = average level of implementation index across all teachers

$\gamma_1 - \gamma_5$ = effect of CSR model key on teacher-level average of implementation index (compared to the omitted category)

$\gamma_6 - \gamma_{11}$ = effect of CSR program on teacher-level average of implementation index (compared to the omitted category)

$\gamma_{12} - \gamma_{18}$ = effect of school-level context variables on teacher-level average of implementation index

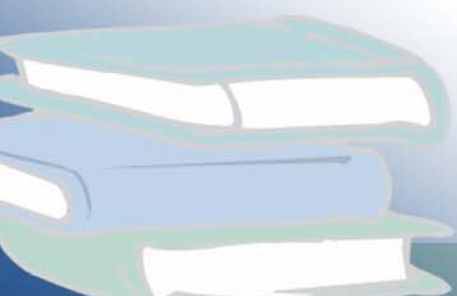
γ_{19} = effect of new principal on teacher-level average of implementation index

γ_{20} = effect of dummy for 3–5 years of CSR model implementation on teacher-level average of implementation index

γ_{21} = effect of dummy for more than 5 years of CSR model implementation on teacher-level average of implementation index

We can summarize the above concisely as follows:

$$Y_{ij} = \beta_{0j} + \sum B * X + \varepsilon_{ij} + \gamma_j$$



Where:

Y_{ij} is the implementation index;

the units of analyses are i teachers that are nested within j schools;

a data matrix X contains values for the predictors;

while B is a set of coefficients to be estimated;

errors, both school-level error γ_j and teacher-level error ε_{ij} are normally distributed, with a mean of zero.

This model tests simultaneously many hypotheses that are of interest to us. We are able to see whether significant differences in the level of implementation exist between comparison schools and CSR schools and between schools having different CSR model keys. In addition, this model allows us to examine what factors (such as the phase of implementation, developers' assistance, principals' instructional leadership, usefulness of developers' assistance, school characteristics, etc.) affect the level of CSR model implementation. In this paper, we have included only tables reporting models in which the CSR Model F implementation key has been used as the omitted reference group.⁷

Results

Year 1 of the Study (2002)

Results for year 2002 are surprisingly consistent across implementation indices and confirmed many initial findings based on descriptive analysis. Our results also confirm findings from earlier studies regarding CSR model implementation (see Table 11).

Do CSR Schools Have Significantly Higher Levels of Implementation?

According to our HLM model, significant differences exist between CSR models and their matched comparison schools regarding the level of implementation. Compared with its comparison schools, CSR Model F schools have about 6% lower level of implementation regarding Use of Technology but about 1.6% higher implementation in Use of Assessments. CSR Model A schools have 5% lower implementation of Parent/Community Involvement than comparison schools but about 3% higher implementation of Pedagogy. CSR Model C schools implement at higher levels than their comparison schools regarding Use of Technology (9% higher), Time Scheduled for Instruction (5% higher), and Pedagogy (3% higher). Statistically significant differences were not found for CSR Model B and CSR Model E.

Our descriptive results showed that CSR Model A schools do implement at higher levels regarding Pedagogy, which is confirmed in our HLM analysis. However, the descriptive analysis also showed higher levels of implementation regarding Influence and Use of Assessments. The positive association still exists in the HLM analysis, but it has lost its statistical significance. The descriptive analysis also showed that CSR Model B schools implement at lower levels than comparison schools regarding Shared

⁷ To obtain all CSR model key differences, as well as differences between all CSR models and their matched comparison schools, the same model was run multiple times with a different omitted reference category. All tables are available upon request.



Decision Making and Parent/Community Involvement; although these negative associations remain in the HLM analysis, they have lost statistical significance due to the control variables included.

Most of these results matched our expectations: CSR Model C is well-known for its focus on technology, CSR Model A for its concentration on pedagogy, and CSR Model F for the importance of assessing and re-evaluating students periodically.

Do Significant Differences Exist Among CSR Model Keys?

The HLM confirms our descriptive results. CSR Model E and CSR Model F model keys are consistently implemented at moderate to high levels; the only exceptions for CSR Model F are Parent/Community Involvement and Use of Assessments. Compared to CSR Model F, both CSR Model B and CSR Model A model keys have higher levels of implementation of Parent/Community Involvement (7% and 11%, respectively), while CSR Model B, CSR Model C, and CSR Model E model keys all have higher levels of implementation regarding Use of Assessments.

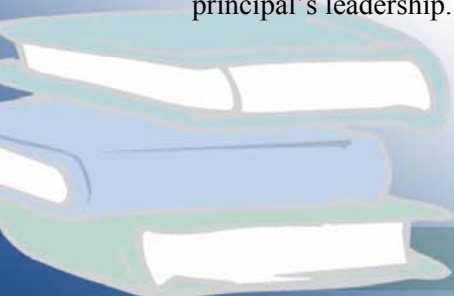
The CSR Model B model key clearly has the lowest levels of implementation regarding Shared Decision Making, Use of Technology, and Inclusion. Compared to the CSR Model F model key, CSR Model B model key has 28% lower implementation of Shared Decision Making, 16.1% lower implementation of Use of Technology in Classrooms, and 52.0% lower implementation of Inclusion. Compared to CSR Model F, the CSR Model A model key also has low levels of implementation regarding Inclusion (29%), Curriculum (8%), Engagement in Informal PD (13%), and Pedagogy (27%). The CSR Model C model key has as its lowest implementation levels Student Grouping and Time Scheduled for Instruction (10% and 36%, respectively, lower than the CSR Model F model key).

These results confirm the idea that some of the model keys are more difficult to implement, especially because few significant differences exist among CSR model schools and comparison schools. Moreover, significant differences between CSR schools and comparison schools tend to exist for implementation indices for which the CSR model key has lower values. These results imply that some CSR models require practices that are not common in all schools. When this is the case, CSR schools tend to have higher levels of implementation than their comparison schools.

Teacher-Level Contextual Variables and Teachers' Background Variables

The results clearly show that teachers' reports of Teacher Community, Principal's Instructional Leadership, and Usefulness of Developer's Assistance are positively related to the level of implementation. The measure of Teacher Community is positively related to many implementation indices: the more often teachers think that they have common goals, the higher the level of implementation of Shared Decision Making; Parent/Community Involvement; Curriculum; Engagement in, and Emphasis of, and Type of PD; and Use of Assessments. The effect of Teacher Community, however, remains small, ranging from 0.5% to approximately 2%. Although the effects are small, they are very consistent, showing the importance of teachers' communicating and sharing similar goals—that is, having a functional professional community.

Similarly, teachers' assessment of Principal's Instructional Leadership is positively related to most implementation indices; the exceptions are Inclusion, Student Grouping, and Curriculum, where no significant relationships exist. The effect of Principal's Instructional Leadership on the level of implementation ranges from 0.5% (Use of Assessments) to 4% (Engagement in Informal PD). The effects are not large but are very consistent, adding to the growing body of research regarding importance of the principal's leadership.



Finally, Usefulness of Developer's Assistance is also positively related to many implementation indices, including Use of Technology, Student Grouping, Parent/Community Involvement, Engagement in Informal PD, Emphasis of and Type of PD, and Use of Assessments. The only significant negative relationship exists between the Usefulness of Developer's Assistance and the level of implementation of Pedagogy. Again, the effect of Developer's Assistance by itself is small (from 0.5% to 2%). However, when effects of Teacher Community, Principal's Instructional Leadership, and Developer's Assistance are combined, the implementation of indices related to PD is boosted by approximately 6%, while Shared Decision Making and Parent/Community Involvement gain about 5%.

The only implementation index that is not affected by Teacher Community, Principal's Instructional Leadership, or Developer's Assistance is Inclusion, which measures how non-English-speaking students and students with disabilities are mainstreamed in regular classrooms. It is possible that decisions regarding mainstreaming students are made at the district level. Our unconditional three-level HLM supports show that 27.5% of variance in Inclusion is located at the district level.

In addition, Teacher's Support for adopting a CSR model is positively related to most implementation indices but is significant only concerning Shared Decision Making and Emphasis of and Type of PD. The implementation level of these indices is about 2% higher if teachers supported the adoption of the CSR model. Although most of the results are not statistically significant, they offer support to the idea that the adoption process is important, but perhaps not as important as having a functional professional community among teachers.

Among the teachers' background variables, being an English teacher clearly stands out. Being an English teacher (rather than a mathematics teacher) is positively associated with the implementation of Pedagogy, Influence of Assessments, Emphasis of and Type of PD, Use of Technology in Classrooms, and Curriculum. However, English teachers are also less faithful in implementing Shared Decision Making, and Engagement in Informal PD. The sizes of the effects, ranging from 1% to 11%, are larger for classroom-related indices such as Curriculum and Pedagogy.

Quite surprisingly, teachers' gender is related to Use of Technology in Classrooms, Parent/Community Involvement, and Use of Assessments. The male teacher variable is positively related to Parent/Community Involvement and Use of Assessments but negatively associated with Use of Technology. Furthermore, being a new teacher at the school has a significant negative relationship to three indices: Use of Technology, Time Scheduled for Instruction, and Use of Assessments.

School-Level Contextual Variables

The variable of schools having middle grades is very consistently and negatively related to 7 of 12 implementation indices. This negative relationship implies that CSR models are difficult to implement with fidelity when schools include a larger range of grades. This result could be caused by added complexity due to curriculum, instructional, and testing requirements that vary by grade level. The effect sizes range from 1 to 5%. For example, the level of Parent/Community Involvement is 5% lower if a school has middle grades.

The size of a school (measured as student enrollment), which is often thought to be an important predictor for the level of implementation, is a significant predictor for only two implementation indices. School size negatively affects implementation of Time Scheduled for Instruction but positively affects Student Grouping. The Percent of Free/Reduced-Price Lunch variable positively affects the fidelity of implementation concerning Shared Decision Making, Use of Technology, Engagement in Informal PD,



and the Emphasis and Type of PD. Schools within high-poverty population areas may receive extra funding and resources for PD and technology; these factors could explain the positive relationships. However, as expected, poverty is negatively related to the implementation of Parent/Community Involvement and Influence of Assessments.

The percentage of non-English-speaking students negatively affects the implementation of Curriculum and Use of Assessments. To better match local needs, schools with large percentages of non-English-speaking students may be modifying the Curriculum as well as Use of Assessments promoted by CSR developers. Schools that did not meet AYP marks tend to implement with less fidelity the PD indices than schools that do make the cut. Although AYP status is negatively associated with the level of many implementation indices, these relationships are not statistically significant.

It is often assumed that change in leadership, such as having a new principal in the school, negatively affects the level of CSR implementation. We included a dummy variable that identifies principals who have less than 1 year of experience in their current schools. This variable is negatively associated only with the level of Curriculum implementation. Thus, although confirming the importance of Principal's Instructional Leadership, our results do not find a substantial effect for instability in leadership.

Finally, whether a school has a CSR coordinator positively affects the implementation of the PD indices and Student Grouping but negatively affects the level of Parent/Community Involvement. Clearly, the presence of a CSR coordinator may help teachers to initiate informal professional networks and to identify PD opportunities that match the CSR model developer's ideals. The negative association between the presence of a CSR coordinator and Parent/Community Involvement is also understandable, because the Parent/Community Involvement index is based on *teachers'* activities that relate to the community and parents. It is possible that most CSR coordinators are very active regarding Parent/Community Involvement; as a result, teachers may not need to be as active as they would be in the absence of a CSR coordinator.

Phase of Implementation

Our preliminary descriptive analysis showed few differences in the level of implementation between schools that have been implementing for 0–2 years, 3–5 years, or more than 5 years. The HLM analysis reconfirms the descriptive results. Only two outcomes, Influence of Assessments and Pedagogy, are affected by the phase of implementation; for both indices, implementation for 5 or more years is related to a higher level of implementation.



Table 11. Regression Results for Level of Implementation, Year 2002

	Shared Decision Making			Use of Technology in Classrooms			Inclusion		
Intercept	0.7079	0.0140	***	0.6117	0.0229	***	0.8205	0.0161	***
CSR Model A key	-0.0475	0.0161	***	-0.0741	0.0263	***	-0.2915	0.0184	***
CSR Model C key	-0.0926	0.0180	***	-0.0556	0.0290	*	-0.0660	0.0201	***
CSR Model D/CSR Model G combined key	0.0001	0.0210		-0.1885	0.0422	***	-0.6718	0.0982	***
CSR Model E key	-0.0910	0.0251	***	0.0000			0.0000		
CSR Model B key	-0.2869	0.0208	***	-0.1610	0.0346	***	-0.5207	0.0236	***
CSR Model F	0.0097	0.0163		-0.0603	0.0268	**	-0.0081	0.0191	
CSR Model A	0.0110	0.0219		-0.0140	0.0357		0.0228	0.0252	
CSR Model C	0.0244	0.0235		0.0905	0.0382	**	0.0361	0.0266	
CSR Model D/CSR Model G Combined	-0.0526	0.0276	*	0.1241	0.0529	**	0.0308	0.1026	
CSR Model E	0.0195	0.0355		0.0000			0.0000		
CSR Model B	-0.0261	0.0348		0.1071	0.0564		0.0158	0.0390	
School size	0.0064	0.0055		-0.0061	0.0094		0.0102	0.0067	
Challenging environment	0.0006	0.0041		-0.0053	0.0070		0.0009	0.0050	
Percent free/reduced-price lunch	0.0123	0.0041	***	0.0153	0.0070	**	-0.0007	0.0050	
Percent non-English-speaking	-0.0057	0.0044		-0.0110	0.0073		-0.0045	0.0051	
School has CSR coordinator	0.0063	0.0088		-0.0015	0.0149		-0.0001	0.0106	
Principal recently joined school	-0.0012	0.0100		-0.0066	0.0172		-0.0066	0.0125	
School has middle grades	-0.0361	0.0096	***	-0.0362	0.0166	**	-0.0145	0.0122	
AYP status 2002–2003	0.0064	0.0089		-0.0199	0.0153		-0.0016	0.0110	
3–5 years of implementation	-0.0139	0.0135		0.0229	0.0225		-0.0073	0.0166	
5 or more years of implementation	0.0009	0.0169		0.0344	0.0279		0.0026	0.0198	
Teacher community: common goals	0.0126	0.0039	***	0.0028	0.0063		0.0025	0.0044	
Principal's instructional leadership	0.0358	0.0040	***	0.0166	0.0064	***	0.0038	0.0044	
Usefulness of developer's assistance	0.0052	0.0036		0.0174	0.0057	***	-0.0030	0.0040	
Teacher recently joined school	-0.0016	0.0071		-0.0464	0.0114	***	-0.0028	0.0079	
English teacher	-0.0195	0.0063	***	0.0438	0.0103	***	0.0591	0.0070	***
Male teacher	0.0047	0.0082		-0.0270	0.0132	*	0.0098	0.0091	
Teacher supported CSR adoption	0.0177	0.0083	**	0.0101	0.0133		-0.0078	0.0092	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 11. Regression Results for Level of Implementation, Year 2002 (continued)

	Student Grouping			Parent/Community Involvement			Curriculum		
Intercept	0.8942	0.0157	***	0.5679	0.0163	***	0.8131	0.0126	***
CSR Model A key	-0.0215	0.0169		0.0725	0.0189	***	-0.0829	0.0134	***
CSR Model C key	-0.1073	0.0193	***	0.0112	0.0206		-0.0121	0.0141	
CSR Model D/CSR Model G combined key	-0.0529	0.0217	**	0.1248	0.0680	*	0.0057	0.0438	
CSR Model E key	-0.0293	0.0259		0.0000			0.0427	0.0251	*
CSR Model B key	-0.0024	0.0225		0.1107	0.0242	***	0.0000		
CSR Model F	-0.0233	0.0170		-0.0098	0.0195		-0.0033	0.0149	
CSR Model A	-0.0096	0.0231		-0.0484	0.0257	*	-0.0175	0.0162	
CSR Model C	0.0054	0.0251		-0.0097	0.0272		0.0165	0.0167	
CSR Model D/CSR Model G combined	0.0445	0.0283		0.0134	0.0736		0.0089	0.0457	
CSR Model E	0.0056	0.0360		0.0000			-0.0362	0.0317	
CSR Model B	0.0460	0.0379		-0.0529	0.0399		0.0000		
School size	0.0118	0.0050	**	0.0013	0.0069		-0.0075	0.0046	
Challenging environment	-0.0049	0.0044		-0.0101	0.0051	**	0.0037	0.0037	
Percent free/reduced-price lunch	-0.0022	0.0043		-0.0127	0.0051	**	0.0015	0.0037	
Percent non-English-speaking	-0.0051	0.0046		0.0066	0.0052		-0.0049	0.0041	
School has CSR coordinator	0.0298	0.0092	***	-0.0226	0.0108	**	-0.0069	0.0076	
Principal recently joined school	-0.0054	0.0103		0.0060	0.0127		-0.0164	0.0089	*
School has middle grades	-0.0214	0.0100	**	-0.0493	0.0125	***	-0.0186	0.0087	**
AYP status 2002–2003	-0.0069	0.0093		-0.0102	0.0112		-0.0126	0.0077	
3–5 years of implementation	-0.0103	0.0139		0.0016	0.0169		0.0153	0.0122	
5 or more years of implementation	0.0217	0.0180		0.0210	0.0201		0.0109	0.0152	
Teacher community: common goals	0.0051	0.0050		0.0161	0.0044	***	0.0065	0.0039	*
Principal's instructional leadership	0.0057	0.0049		0.0172	0.0044	***	0.0051	0.0039	
Usefulness of developer's assistance	0.0091	0.0045	**	0.0198	0.0039	***	-0.0004	0.0035	
Teacher recently joined school	0.0030	0.0089		-0.0050	0.0078		-0.0003	0.0071	
English teacher	-0.0132	0.0081		-0.0060	0.0070		0.1144	0.0086	***
Male teacher	0.0091	0.0105		0.0215	0.0090	**	0.0025	0.0082	
Teacher supported CSR adoption	-0.0063	0.0105		0.0123	0.0092		-0.0023	0.0083	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 11. Regression Results for Level of Implementation, Year 2002 (continued)

	Time Scheduled for Instruction		Engagement in Informal PD				Influence of Assessments		
Intercept	0.9051	0.0110	***	0.7723	0.0147	***	0.7461	0.0109	***
CSR Model A key	0.0391	0.0126	***	-0.1336	0.0162	***	-0.0400	0.0119	***
CSR Model C key	-0.3672	0.0141	***	-0.0902	0.0180	***	0.0020	0.0132	
CSR Model D/CSR Model G combined key	-0.1406	0.0164	***	0.0588	0.0624		0.0423	0.0463	
CSR Model E key	0.0147	0.0195		-0.0375	0.0248		0.1105	0.0181	***
CSR Model B key	-0.0204	0.0163		0.0213	0.0207		0.0176	0.0154	
CSR Model F	-0.0163	0.0127		-0.0072	0.0167		-0.0100	0.0122	
CSR Model A	-0.0182	0.0171		0.0161	0.0222		0.0173	0.0163	
CSR Model C	0.0521	0.0184	***	0.0160	0.0236		0.0076	0.0173	
CSR Model D/CSR Model G combined	0.0572	0.0214	***	-0.0786	0.0667		-0.0310	0.0492	
CSR Model E	-0.0299	0.0275		0.0383	0.0352		-0.0371	0.0256	
CSR Model B	0.0114	0.0272		0.0112	0.0352		0.0018	0.0260	
School size	-0.0129	0.0042	***	0.0002	0.0053		-0.0007	0.0038	
Challenging environment	0.0002	0.0032		0.0062	0.0043		-0.0040	0.0031	
Percent free/reduced-price lunch	0.0004	0.0032		0.0073	0.0042	*	-0.0085	0.0031	***
Percent non-English-speaking	0.0029	0.0034		-0.0048	0.0044		-0.0013	0.0032	
School has CSR coordinator	-0.0018	0.0068		0.0207	0.0090	*	0.0049	0.0066	
Principal recently joined school	0.0059	0.0077		-0.0036	0.0104		-0.0037	0.0076	
School has middle grades	-0.0123	0.0074	*	-0.0095	0.0102		-0.0190	0.0075	**
AYP status 2002–2003	0.0021	0.0069		-0.0258	0.0092	***	0.0061	0.0067	
3–5 years of implementation	0.0138	0.0105		-0.0176	0.0144		0.0005	0.0105	
5 or more years of implementation	0.0097	0.0133		0.0078	0.0173		0.0223	0.0127	*
Teacher community: common goals	-0.0036	0.0032		0.0107	0.0044	**	-0.0002	0.0034	
Principal's instructional leadership	0.0079	0.0032	**	0.0443	0.0044	***	0.0167	0.0033	***
Usefulness of developer's assistance	-0.0009	0.0029		0.0100	0.0040	**	0.0033	0.0030	
Teacher recently joined school	-0.0102	0.0057	*	-0.0108	0.0079		0.0008	0.0061	
English teacher	-0.0045	0.0051		-0.0186	0.0071	***	0.0171	0.0055	***
Male teacher	0.0079	0.0067		0.0083	0.0092		-0.0058	0.0071	
Teacher supported CSR adoption	0.0071	0.0067		0.0096	0.0092		0.0105	0.0071	

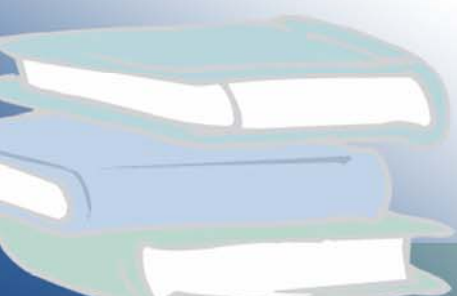
Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 11. Regression Results for Level of Implementation, Year 2002 (continued)

	Use of Assessments			Pedagogy			Emphasis of and Type of PD		
Intercept	0.8456	0.0081	***	0.9389	0.0096	***	0.7141	0.0152	***
CSR Model A key	-0.0061	0.0091		-0.2668	0.0103	***	-0.0151	0.0168	
CSR Model C key	0.0578	0.0101	***	-0.0595	0.0110	***	0.0511	0.0189	***
CSR Model D/CSR Model G combined key	0.0293	0.0117	**	-0.0167	0.0353		0.0056	0.0216	
CSR Model E key	0.0428	0.0140	***	-0.0399	0.0142	***	0.0205	0.0258	
CSR Model B key	0.0523	0.0117	***	-0.0434	0.0121	***	-0.0318	0.0216	
CSR Model F	0.0167	0.0091	*	0.0055	0.0115		-0.0138	0.0169	
CSR Model A	0.0209	0.0123		0.0280	0.0125	**	0.0097	0.0228	
CSR Model C	0.0069	0.0132		0.0294	0.0129	**	-0.0041	0.0246	
CSR Model D/CSR Model G combined	0.0359	0.0153	**	-0.0025	0.0368		-0.0157	0.0282	
CSR Model E	-0.0006	0.0196		-0.0072	0.0192		-0.0189	0.0360	
CSR Model B	0.0054	0.0197		-0.0107	0.0197		-0.0011	0.0369	
School size	-0.0002	0.0029		-0.0037	0.0034		0.0051	0.0053	
Challenging environment	0.0045	0.0023	*	-0.0008	0.0027		0.0007	0.0043	
Percent free/reduced-price lunch	0.0003	0.0023		0.0009	0.0027		0.0162	0.0043	***
Percent non-English-speaking	-0.0046	0.0025	*	0.0013	0.0030	*	-0.0068	0.0045	
School has CSR coordinator	-0.0003	0.0049		0.0029	0.0056		0.0201	0.0091	**
Principal recently joined school	-0.0009	0.0055		0.0027	0.0064		-0.0157	0.0103	
School has middle grades	-0.0042	0.0053		-0.0085	0.0061		0.0053	0.0099	
AYP status 2002–2003	-0.0018	0.0050		-0.0027	0.0056		-0.0235	0.0092	**
3–5 years of implementation	-0.0103	0.0075		-0.0106	0.0091		-0.0097	0.0139	
5 or more years of implementation	-0.0055	0.0095		0.0009	0.0114	*	0.0087	0.0178	
Teacher community: common goals	0.0045	0.0024	*	0.0003	0.0030		0.0177	0.0047	***
Principal's instructional leadership	0.0064	0.0024	***	0.0060	0.0029	*	0.0200	0.0046	***
Usefulness of developer's assistance	0.0058	0.0022	***	-0.0045	0.0027	*	0.0260	0.0042	***
Teacher recently joined school	-0.0072	0.0043	*	-0.0043	0.0054		-0.0076	0.0084	
English teacher	-0.0038	0.0039		0.0370	0.0060	***	0.0395	0.0076	***
Male teacher	0.0110	0.0051	*	-0.0012	0.0062		-0.0090	0.0098	
Teacher supported CSR adoption	0.0063	0.0051		0.0060	0.0064		0.0181	0.0098	*

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Year 3 of the Study (2004)

The statistical model applied for year 3 (2004) data was similar to the model used for year 1 (2002) data, with one exception. The variable about teachers' support for CSR adoption was dropped because significantly more teachers were not in the school when the program was adopted, making the measure less informative than in 2002. Because of teacher-level attrition, the sample in 2004 was also smaller than in 2002.

Significant differences exist between CSR schools and their matched comparison schools (see Table 12). CSR Model F schools have a 3% higher level of implementation regarding Engagement in Informal PD. CSR Model A schools do better than comparison schools regarding Shared Decision Making (8% higher), Use of Technology in Classrooms (7% higher), and Influence of Assessments (3% higher). CSR Model C schools implement Parent/Community Involvement and Curriculum at higher levels (7% and 5%, respectively), while doing worse than their comparison schools regarding Pedagogy (6% lower). CSR Model B schools implement Use of Technology better than their comparison schools (11% higher) but lag behind regarding Pedagogy implementation (5% lower). The biggest change is for CSR Model E schools; these are significantly lower implementers than their matched comparison schools in Engagement in Informal PD (10%), Influence of Assessments (11%), and Pedagogy (6%).

Large differences in the level of implementation exist between different CSR model keys. As in 2002, CSR Model F and CSR Model E model keys continue to be highly implemented compared to other CSR program keys. CSR Model B model key continues to be implemented at a low level regarding Shared Decision Making and Inclusion. Similarly, the CSR Model A model key continues to be especially difficult to implement regarding Pedagogy, Engagement in Informal PD, and Shared Decision Making. The CSR Model C model key still distinguishes itself from other CSR model keys by a lower level of implementation regarding Time Scheduled for Instruction.

From the teacher-level contextual variables, Principal's Instructional Leadership and Usefulness of Developer's Assistance continue to have positive effects on implementation across several implementation indices. Being an English teacher is still very consistently and positively related to implementation for most implementation indices (an exception being the Student Grouping index). Being a new teacher in the school is negatively related to implementation of Engagement in Informal PD, Use and Influence of Assessments indices, and Curriculum. As before, being a male teacher is positively related to implementation of Use of Assessments, Parent/Community Involvement, Student Grouping, and Engagement in Informal PD.

The largest change regarding school-level contextual variables is that schools with middle grades do not consistently have lower levels of implementation. School size is often related to level of implementation but not consistently across different implementation indices. School size is positively related to Shared Decision Making, Student Grouping, and Emphasis and Type of PD. However, the effect of school size is negative concerning Time Scheduled for Instruction, Use of Assessments, and Curriculum.

The Percentage of Students Receiving Free/Reduced-Price Lunch continues to be positively related to engagement in Informal PD, Emphasis and Type of PD, Pedagogy, and Use of Technology in Classrooms. Again, it is possible that schools with higher levels of poverty receive additional resources



for PD and technology. The percentage of non-English-speaking students is negatively related to Inclusion (schools with larger proportions of non-English-speaking students experience more difficulty in mainstreaming their students), Pedagogy, and Emphasis and Type of PD. Quite surprisingly, this variable was positively related to Parent/Community Involvement in 2004. Schools' AYP status continues to be negatively associated with most implementation indices and significantly affects the level of implementation concerning Influence of Assessments, Pedagogy, and Emphasis of and Type of PD.

As in 2002, the phase of implementation is rarely a significant predictor of the level of implementation. Schools that have implemented a CSR program for more than 5 years have a lower level of implementation regarding Use of Assessments than schools that have been implementing for less than 2 years. However, schools that have been implementing more than 5 years have a higher level of implementation of Student Grouping.

The lesson learned from the cross-sectional analysis is as follows: systematic differences between CSR schools and comparison schools do not exist, while the CSR model implementation keys do influence the level of implementation in a consistent fashion. In addition, most contextual variables do not consistently predict the level of implementation for a good reason: it makes sense that different dimensions of CSR implementation are related to different contextual factors. This complex relationship is what makes implementing a CSR model a challenge: some components of CSR models are likely to be challenging in any school. This is also what makes our findings related to principal's leadership, teacher community, and developer's assistance valuable: these variables predict higher level of implementation across different components of CSR models.



Table 12. Regression Results for Level of Implementation, Year 2004

	Shared Decision Making		Use of Technology in Classrooms			Inclusion			
Intercept	0.6389	0.0203	***	0.5997	0.0330	***	0.7942	0.0288	***
CSR Model A key	-0.1052	0.0181	***	-0.0525	0.0279	*	-0.0956	0.0250	***
CSR Model C key	-0.0215	0.0238		-0.0493	0.0369		-0.0176	0.0324	
CSR Model D/CSR Model G combined key	0.1060	0.0265	***	-0.1109	0.0409	***	-0.3851	0.0769	***
CSR Model E key	0.0210	0.0334		0.0000			0.0000		
CSR Model B key	-0.0733	0.0269	***	-0.1390	0.0420	***	-0.3919	0.0383	***
CSR Model F	0.0171	0.0189		-0.0130	0.0294		-0.0032	0.0256	
CSR Model A	0.0805	0.0266	***	0.0692	0.0411	*	-0.0228	0.0385	
CSR Model C	0.0116	0.0360		0.0566	0.0567		-0.0689	0.0473	
CSR Model D/CSR Model G combined	-0.1647	0.0378	***	0.0623	0.0584		-0.1333	0.0880	
CSR Model E	-0.0015	0.0481		0.0000			0.0000		
CSR Model B	-0.0068	0.0416		0.1129	0.0649	*	0.1293	0.0578	
School size	0.0111	0.0058	*	-0.0135	0.0088		-0.0015	0.0081	
Challenging environment	-0.0052	0.0057		-0.0045	0.0091		-0.0138	0.0083	*
Percent free/reduced-price lunch	0.0087	0.0054		0.0128	0.0090	**	0.0067	0.0085	
Percent non-English-speaking	0.0057	0.0058		-0.0058	0.0090		-0.0327	0.0080	***
School has CSR coordinator	-0.0185	0.0144		0.0004	0.0227		0.0031	0.0206	
Principal recently joined school	-0.0034	0.0151		-0.0418	0.0240	*	-0.0023	0.0212	
School has middle grades	0.0176	0.0126		-0.0143	0.0201		0.0206	0.0185	
AYP status 2002–2003	-0.0051	0.0122		0.0014	0.0195		0.0067	0.0174	
3–5 years of implementation	-0.0231	0.0184		-0.0157	0.0287		0.0279	0.0267	
5 or more years of implementation	-0.0290	0.0197		-0.0309	0.0320		0.0173	0.0273	
Teacher community: common goals	-0.0040	0.0058		0.0073	0.0098		-0.0041	0.0083	
Principal's instructional leadership	0.0358	0.0061	***	0.0209	0.0101	*	0.0049	0.0088	
Usefulness of developer's assistance	0.0116	0.0048	***	0.0196	0.0081	**	-0.0099	0.0068	
Teacher recently joined school	-0.0309	0.0300		-0.0075	0.0496		0.1088	0.0392	***
English teacher	0.0183	0.0092	*	0.0498	0.0151	***	0.0244	0.0125	***
Male teacher	0.0141	0.0117		-0.0243	0.0194		0.0145	0.0159	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 12. Regression Results for Level of Implementation, Year 2004 (continued)

	Student Grouping		Parent/Community Involvement			Curriculum			
Intercept	0.9290	0.0204	***	0.6238	0.0244	***	0.7850	0.0188	***
CSR Model A key	-0.0541	0.0170	***	0.0336	0.0210		X-0.0152	0.0165	
CSR Model C key	-0.0792	0.0238	***	0.0157	0.0271		0.0069	0.0202	
CSR Model D/CSR Model G combined key	-0.0510	0.0254	*	0.0395	0.0623		-0.0279	0.0397	
CSR Model E key	-0.0201	0.0316		0.0000			0.0638	0.0434	
CSR Model B key	-0.0077	0.0260		0.0528	0.0309	*	0.0000		
CSR Model F	-0.0240	0.0184		-0.0120	0.0220		0.0177	0.0188	
CSR Model A	0.0235	0.0258		-0.0135	0.0311		-0.0054	0.0210	
CSR Model C	-0.0129	0.0362		0.0718	0.0421	*	0.0528	0.0288	*
CSR Model D/CSR Model G combined	0.0626	0.0368	*	0.1070	0.0727		0.0565	0.0442	
CSR Model E	-0.0106	0.0441		0.0000			0.0211	0.0502	
CSR Model B	-0.0100	0.0413		0.0422	0.0479		0.0000		
School size	0.0048	0.0048	**	0.0038	0.0070		-0.0144	0.0053	*
Challenging environment	0.0030	0.0055		-0.0014	0.0070		-0.0069	0.0053	
Percent free/reduced-price lunch	-0.0031	0.0053		-0.0042	0.0072		0.0019	0.0055	
Percent non-English-speaking	0.0028	0.0055		0.0161	0.0069	*	0.0085	0.0053	
School has CSR coordinator	-0.0060	0.0140		-0.0189	0.0173		0.0169	0.0134	
Principal recently joined school	0.0233	0.0150		0.0151	0.0176		0.0105	0.0143	
School has middle grades	-0.0170	0.0121		-0.0370	0.0156	*	-0.0067	0.0119	
AYP status 2002–2003	-0.0067	0.0119		0.0014	0.0148		-0.0158	0.0114	
3–5 years of implementation	-0.0016	0.0177		-0.0284	0.0232		-0.0282	0.0180	
5 or more years of implementation	0.0594	0.0195	***	-0.0296	0.0236		0.0119	0.0202	
Teacher community: common goals	-0.0005	0.0065		0.0141	0.0069	*	-0.0077	0.0062	
Principal's instructional leadership	0.0051	0.0065		0.0078	0.0072		0.0089	0.0063	
Usefulness of developer's assistance	0.0091	0.0052	*	0.0248	0.0058	***	0.0064	0.0050	
Teacher recently joined school	-0.0044	0.0324		-0.0372	0.0357		-0.0516	0.0311	*
English teacher	-0.0426	0.0101	***	0.0069	0.0107		0.0830	0.0125	***
Male teacher	0.0318	0.0129	*	0.0253	0.0138	*	-0.0096	0.0120	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Table 12. Regression Results for Level of Implementation, Year 2004 (continued)

	Time Scheduled for Instruction			Engagement in Informal PD		Influence of Assessments			
Intercept	0.9154	0.0143	***	0.7430	0.0193	***	0.7752	0.0152	***
CSR Model A key	0.0228	0.0125	***	-0.1332	0.0163	***	-0.0808	0.0132	***
CSR Model C key	-0.4568	0.0165	***	-0.0432	0.0223	*	0.0173	0.0173	
CSR Model D/CSR Model G combined key	-0.1160	0.0184	***	-0.0018	0.0473		-0.0310	0.0388	
CSR Model E key	-0.0016	0.0231		0.0750	0.0300	**	0.0773	0.0245	***
CSR Model B key	-0.0046	0.0186		0.0639	0.0243	***	0.0258	0.0198	
CSR Model F	0.0010	0.0132		0.0332	0.0176	*	0.0155	0.0139	
CSR Model A	0.0017	0.0185		0.0336	0.0249		0.0337	0.0197	*
CSR Model C	0.0210	0.0250		0.0259	0.0340		0.0030	0.0268	
CSR Model D/CSR Model G combined	0.0473	0.0262	*	-0.0848	0.0562		0.0451	0.0455	
CSR Model E	0.0284	0.0333		-0.0991	0.0417	*	-0.1125	0.0352	*
CSR Model B	-0.0203	0.0288		-0.0448	0.0389		0.0378	0.0307	
School size	-0.0096	0.0040	**	-0.0078	0.0047		-0.0004	0.0042	
Challenging environment	0.0035	0.0040		-0.0050	0.0053		0.0022	0.0043	
Percent free/reduced-price lunch	0.0002	0.0038		0.0141	0.0054	***	0.0031	0.0043	
Percent non-English-speaking	-0.0036	0.0040		-0.0021	0.0053		-0.0002	0.0043	
School has CSR coordinator	-0.0054	0.0100		0.0067	0.0136	*	0.0014	0.0108	
Principal recently joined school	0.0186	0.0105	*	-0.0143	0.0141		-0.0053	0.0111	
School has middle grades	-0.0164	0.0087	*	0.0086	0.0119		0.0097	0.0096	
AYP status 2002–2003	0.0064	0.0085		-0.0034	0.0116		-0.0224	0.0092	*
3–5 years of implementation	0.0100	0.0128		0.0100	0.0183		-0.0149	0.0145	
5 or more years of implementation	-0.0087	0.0138		-0.0057	0.0185		-0.0091	0.0146	
Teacher community: common goals	0.0091	0.0041	**	0.0105	0.0061	*	0.0053	0.0044	
Principal’s instructional leadership	0.0008	0.0042		0.0350	0.0062	***	0.0027	0.0046	
Usefulness of developer’s assistance	0.0049	0.0033		0.0183	0.0050	***	0.0092	0.0037	**
Teacher recently joined school	0.0074	0.0208		-0.0928	0.0310	***	-0.0417	0.0233	*
English teacher	-0.0055	0.0064		0.0292	0.0096	***	0.0328	0.0070	***
Male teacher	0.0111	0.0082		0.0215	0.0121	*	-0.0185	0.0090	*

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 12. Regression Results for Level of Implementation, Year 2004 (continued)

	Use of Assessments			Pedagogy			Emphasis and Type of PD		
Intercept	0.8424	0.0110	***	0.9570	0.0126	***	0.7230	0.0216	***
CSR Model A key	-0.0331	0.0092	***	-0.3295	0.0114	***	0.0179	0.0187	
CSR Model C key	0.0577	0.0128	***	-0.0616	0.0142	***	0.0075	0.0253	
CSR Model D/CSR Model G combined key	0.0308	0.0136	**	-0.0969	0.0288	***	-0.0073	0.0276	
CSR Model E key	0.0700	0.0177	***	-0.0254	0.0182		0.0551	0.0346	
CSR Model B key	0.0470	0.0140	***	-0.0593	0.0151	***	-0.0053	0.0284	
CSR Model F	0.0046	0.0098		-0.0004	0.0132		-0.0241	0.0197	
CSR Model A	0.0202	0.0140		0.0075	0.0149		0.0141	0.0279	
CSR Model C	-0.0177	0.0193		-0.0424	0.0201	**	0.0215	0.0381	
CSR Model D/CSR Model G combined	-0.0052	0.0199		0.0162	0.0324		-0.0329	0.0398	
CSR Model E	-0.0255	0.0242		-0.0576	0.0245	**	-0.0804	0.0492	
CSR Model B	0.0082	0.0223		-0.0484	0.0226	**	0.0171	0.0443	
School size	-0.0094	0.0026	***	-0.0016	0.0038		0.0135	0.0056	**
Challenging environment	0.0039	0.0030		-0.0010	0.0036		-0.0019	0.0060	
Percent free/reduced-price lunch	0.0049	0.0029		0.0078	0.0035	**	0.0175	0.0057	***
Percent non-English-speaking	0.0025	0.0030	*	-0.0070	0.0036	*	-0.0123	0.0060	*
School has CSR coordinator	-0.0075	0.0075		0.0115	0.0088		0.0131	0.0151	
Principal recently joined school	0.0051	0.0081		-0.0072	0.0093		0.0020	0.0160	
School has middle grades	0.0094	0.0065		-0.0040	0.0077		-0.0044	0.0131	
AYP status 2002–2003	-0.0028	0.0064		-0.0128	0.0075	*	-0.0316	0.0128	**
3–5 years of implementation	0.0118	0.0095		0.0079	0.0122		0.0036	0.0192	
5 or more years of implementation	-0.0179	0.0105	*	0.0137	0.0133		0.0076	0.0208	
Teacher community: common goals	0.0013	0.0034		0.0033	0.0039		0.0011	0.0065	
Principal's instructional leadership	0.0028	0.0035		-0.0053	0.0040		0.0447	0.0068	***
Usefulness of developer's assistance	0.0064	0.0028	**	-0.0015	0.0032		0.0193	0.0054	***
Teacher recently joined school	-0.0263	0.0179		-0.0047	0.0206		-0.0421	0.0334	
English teacher	0.0168	0.0054	***	0.0937	0.0075	***	0.0548	0.0103	***
Male teacher	0.0115	0.0070	*	-0.0013	0.0077		-0.0031	0.0132	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



What Predicts Change in Implementation Between 2002 and 2004?

Two-level HLM, including teacher and school levels, was conducted to examine what predicts the change in implementation between 2002 and 2004. The outcomes in this analysis are difference scores calculated by subtracting the 2002 implementation index value from the 2004 implementation index value. Our descriptive analysis showed that the values for most implementation indices increased between 2002 and 2004. The exceptions were those indices that already had a high level of implementation in 2002 (Curriculum, and Use and Influence of Assessments). The goal was to find out what factors predict the gains in the implementation indices. We assumed that the following variables predict the teacher-level implementation:

1. Two sets of dummy variables: the first set is for different CSR model keys (identifying CSR schools and their matched comparison schools); the second set is for schools implementing specific CSR models (school-level variables).
2. Two dummy variables measuring level of implementation in year 2002 (medium and high levels of implementation; the low level of implementation is used as an omitted reference category).
3. Variables controlling for school context: Student Enrollment, Percentage of Students Receiving Free/Reduced-Price Lunch, Percentage of Non-English-Speaking Students, Existence of a CSR Coach or Coordinator in the Schools, School with Middle Grades, and School's AYP Status in 2002 (school-level variables).
4. Principal leadership: Principal's Instructional Leadership Scale (average from 2002 and 2004) (teacher-level variable) and change in Instructional Leadership Scale, a dummy variable recognizing new principals with experience of 1 year or less in their current school (school-level variable).
5. Teacher background: dummy variable identifying new teachers with experience of 1 year or less in their current school, dummy variable distinguishing between English and mathematics teachers, Teacher Community Scale (average from 2002 and 2004), and Change in Teacher Community Scale (between 2002 and 2004).
6. Usefulness of Developer's Support Scale (average from 2002 and 2004) and Change in Usefulness of Developer's Support between 2002 and 2004 (teacher-level scale).
7. Years of CSR implementation: Dummy variables for 3–5 years and for more than 5 years of implementation (school-level variables).

The change in implementation may be affected by both the level and change of our predictors. We examined our independent variables to assess whether large changes took place between 2002 and 2004. Because the school-level variables measuring school characteristics did not change significantly between 2002 and 2004, we decided to use 2002 values in our analysis. For instance, we hypothesized that the change in implementation is affected by the level of poverty in 2002, not by the change in level of poverty between 2002 and 2004. However, we decided to use both the level and change for variables that are likely to fluctuate more over time. These variables include teacher-level reports on Teacher Community, Usefulness of Developer's Assistance, and Principal's Instructional Leadership. Thus we assume that the change in these variables, not just their levels, is predicting change in implementation.



We can summarize the above concisely as follows:

$$Y_{ij} = \beta_{0j} + \sum B * X + \varepsilon_{ij} + \gamma_j$$

Where:

Y_{ij} is the difference in implementation index between 2002 and 2004;

the units of analyses are i teachers that are nested within j schools;

a data matrix X contains values for the predictors;

while B is a set of coefficients to be estimated; and

errors, both school-level error γ_j and teacher-level error ε_{ij} are normally distributed with a mean of zero.

As before, this model allows us to address several hypotheses simultaneously. Our main interest is to see whether the level of implementation change differs between CSR model schools and their matched control schools, but we also explore whether CSR model keys are related to change in implementation. Finally, our model will show what other school- and teacher-level factors are related to change in implementation.

Results

Significant differences existed between CSR model schools and their matched comparison schools (see Table 13). CSR Model F schools had a larger gain (approximately 4%) in implementation than control schools in Influence of Assessments. In addition, CSR Model F schools gained 6% more than comparison schools in Curriculum implementation. CSR Model A schools gained about 8% more than control schools concerning Use of Technology, while the control schools made approximately 3% larger gains in Pedagogy. CSR Model C increased its implementation more than control schools in Emphasis and Type of PD (9%), Parent/Community Involvement (9%), and Curriculum (5%). CSR Model E schools gained less than their comparison schools in Engagement in Informal PD (10%). No significant differences existed between CSR Model B schools and their comparison schools.

Significant differences also existed regarding CSR model keys. As in the cross-sectional analysis, we used CSR Model F as our omitted reference category. According to our cross-sectional analysis, CSR Model F and CSR Model E model keys were consistently highly implemented compared to other CSR model keys in regard to most implementation indices. Therefore, we expected that these CSR model keys would not be able to gain as much as other CSR model keys over time.

Our results showed that no significant differences exist among the CSR model keys concerning several implementation indices, including Use of Technology, Student Grouping, Engagement in Informal PD, Use of Assessments, and Pedagogy. However, some differences existed both statistically and substantively. CSR Model B and CSR Model E model keys gained more than CSR Model F regarding



Shared Decision Making (approximately 8% and 17%, respectively), while the CSR Model A model key gained 8% less than the CSR Model F model key in Shared Decision Making. Moreover, CSR Model A and CSR Model B model keys gained 20% and 15%, respectively, more than CSR Model F concerning Inclusion. CSR Model A model key also gained 8% more in Pedagogy, and CSR Model C gained 7% more regarding Engagement in Informal PD. However, the CSR Model F model key still gained significantly more than CSR Model A and CSR Model C model keys regarding Time Scheduled for Instruction (3% and 10%, respectively). Similarly, the CSR Model A and CSR Model E model keys gained 5% and 8% less than CSR Model F concerning Influence of Assessments. Finally, the CSR Model A model key also gained 16% less than CSR Model F in Emphasis and Type of PD.

These results clearly show that the CSR model keys that had very low implementation levels in 2002 significantly increased their level of implementation. In addition, although the CSR Model F model key was already highly implemented in 2002, the level of implementation increased more than some other CSR model keys regarding those implementation indices that are particularly relevant for CSR Model F: Influence of Assessments and Time Scheduled for Instruction. More generally, the level of implementation in 2002 had a negative and significant relationship to the gain in implementation regarding Use of Technology, Engagement in Informal PD, Emphasis of PD, and Influence of Assessments. This finding implies that high starting values are negatively associated with the gain in implementation. That is, if the value of an implementation index was high in 2002, there was little room for additional positive change to take place.

In the cross-sectional analysis, the level of implementation was consistently predicted by Principal's Instructional Leadership, Teacher Community, and Usefulness of CSR Model Developer's Assistance. In this model, we added both an average level and a change variable to see whether either the level in 2002 or the change from 2002 to 2004 predicts change in implementation. Change in Principal's Instructional Leadership was positively related to gain in Engagement in Informal PD, Emphasis of and Type of PD, Shared Decision Making, Inclusion, and Influence of Assessments. However, the level of Principal's Instructional Leadership was negatively related to gain in implementation of Influence of Assessments and Pedagogy, but positively related to Emphasis and Type of PD. The size of the effects varied from 1.4% to approximately 5%.

Usefulness of Developer's Assistance, a strong predictor in both cross-sectional models, lost its significance regarding change in implementation. Significant positive effects still existed regarding the level of Developer's Assistance in 2002 concerning Use of Assessments (higher level of Usefulness of Assistance in 2002 was related to gain in implementation between 2002 and 2004), while change in Usefulness of Developer's Assistance was related to a positive gain in implementation regarding both Parent/Community Involvement and Emphasis and Type of PD.

Our measure of teachers' professional community—Teacher Community—also lost its predictive power in the change model. The level of Teacher Community in 2002 was negatively related to change in implementation regarding Shared Decision Making. The change in Teacher Community was associated with a gain in implementation of Parent/Community Involvement, but the change was negatively related to the gain in Influence of Assessments.



From the teacher-level variables, being an English teacher clearly stands out. In cross-sectional analysis, being an English teacher was related to a lower level of implementation regarding Shared Decision Making and Engagement in Informal PD but it was positively related to the implementation level of Inclusion and Curriculum, among other indices. In the change model, these results were reversed: being an English teacher predicted positive change in implementation of Shared Decision Making and Engagement in Informal PD but negative change in implementation of Inclusion and Curriculum. Clearly, English teachers could improve their implementation regarding Shared Decision Making and Engagement in Informal PD, but they may have faced ceiling effects regarding Inclusion and Curriculum.

None of the school-level control variables showed a systematic relationship to the change in implementation. However, having a CSR coach in 2002 affected negatively two of the implementation indices: Student Grouping and Curriculum. This effect might be related to the high level of implementation of these indices in 2002 in schools having a CSR coordinator.

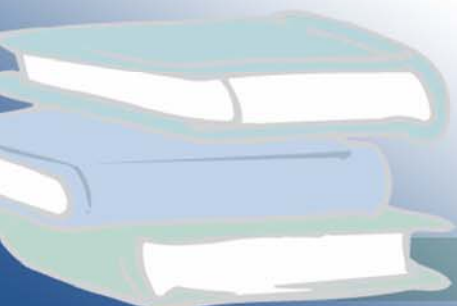


Table 13. Regression Results for Change in Implementation From 2002 to 2004

	Shared Decision Making		Use of Technology in Classrooms				Inclusion		
Intercept	0.0476	0.0643		0.0026	0.0931		-0.0436	0.0869	
CSR Model A key	-0.0822	0.0277	***	-0.0463	0.0371		0.2037	0.0339	***
CSR Model C key	0.0353	0.0311		0.0484	0.0404		0.0446	0.0375	
CSR Model D/CSR Model G combined key	0.0836	0.0358	**	-0.0225	0.0549		0.3804	0.1242	***
CSR Model E key	0.0787	0.0399	**	0.0000			0.0000		
CSR Model B key	0.1715	0.0329	***	0.0200	0.0429		0.1534	0.0409	***
CSR Model F	-0.0290	0.0257		0.0537	0.0346		-0.0150	0.0319	
CSR Model A	0.0367	0.0322		0.0827	0.0425	*	-0.0364	0.0412	
CSR Model C	-0.0067	0.0434		-0.0281	0.0580		-0.0294	0.0524	
CSR Model D/CSR Model G combined	-0.0869	0.0470	*	0.0027	0.0731		-0.1713	0.1415	
CSR Model E	-0.0411	0.0559		0.0000			0.0000		
CSR Model B	0.0329	0.0518		-0.0363	0.0681		0.0262	0.0647	
Medium implementation 2002	-0.0277	0.0210		-0.0729	0.0289	**	0.0215	0.0281	
High implementation 2002	-0.0410	0.0262		-0.0886	0.0356	**	-0.0048	0.0341	
School size	0.0005	0.0014		-0.0016	0.0018		-0.0032	0.0017	*
Percent free/reduced-price lunch	-0.0127	0.0180		-0.0005	0.0250		0.0005	0.0243	
Percent non-English-speaking	0.0156	0.0320		0.0290	0.0421		-0.0571	0.0387	
School has CSR coordinator	0.0157	0.0138		-0.0083	0.0190		0.0180	0.0185	
Principal recently joined school	-0.0259	0.0178		-0.0142	0.0245		0.0086	0.0244	
School has middle grades	0.0080	0.0150		0.0098	0.0208		0.0089	0.0200	
AYP status 2002–2003	-0.0111	0.0146		0.0167	0.0203		0.0183	0.0198	
3–5 years of implementation	0.0064	0.0215		-0.0560	0.0294	*	0.0283	0.0296	
5 or more years of implementation	-0.0162	0.0251		-0.0651	0.0338	*	-0.0042	0.0307	
Teacher community: common goals	-0.0285	0.0163	*	0.0265	0.0237		-0.0143	0.0220	
Change in teacher community	-0.0031	0.0106		0.0057	0.0158		-0.0238	0.0148	
Principal's instructional leadership	0.0066	0.0142		0.0143	0.0204		0.0278	0.0193	
Change in principal's instructional leadership	0.0431	0.0102	***	-0.0141	0.0154		0.0242	0.0145	*
Usefulness of developer's assistance	-0.0010	0.0150		-0.0165	0.0216		0.0039	0.0204	
Change in usefulness of developer's assistance	0.0133	0.0099		0.0234	0.0148		0.0177	0.0136	
Teacher recently joined school	0.0001	0.0132		0.0095	0.0192		-0.0012	0.0178	
English teacher	0.0382	0.0120	***	-0.0012	0.0178		-0.0644	0.0163	***
Male teacher	-0.0025	0.0160		0.0065	0.0238		0.0158	0.0219	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table 13. Regression Results for Change in Implementation From 2002 to 2004 (continued)

	Student Grouping			Parent/Community Involvement			Curriculum		
Intercept	-0.0018	0.0736		0.2152	0.0628	***	0.1225	0.0661	*
CSR Model A key	0.0197	0.0299		-0.0388	0.0251		0.0825	0.0258	***
CSR Model C key	0.0393	0.0335		-0.0147	0.0273		0.0361	0.0244	
CSR Model D/CSR Model G combined key	-0.0015	0.0381		0.0108	0.0563		0.0332	0.0464	
CSR Model E key	0.0332	0.0420		0.0000			0.0664	0.0479	
CSR Model B key	0.0111	0.0360		-0.0177	0.0282		0.0000		
CSR Model F	0.0327	0.0278		0.0297	0.0230		0.0604	0.0250	*
CSR Model A	0.0077	0.0347		0.0197	0.0286		0.0232	0.0238	
CSR Model C	0.0351	0.0480		0.0884	0.0385	**	0.0544	0.0306	*
CSR Model D/CSR Model G combined	0.0521	0.0507		0.0481	0.0689		0.0045	0.0541	
CSR Model E	0.0036	0.0590		0.0000			0.0605	0.0564	
CSR Model B	-0.0319	0.0566		0.0364	0.0458		0.0000		
Medium implementation 2002	0.0170	0.0228		-0.0257	0.0207		0.0124	0.0198	
High implementation 2002	0.0285	0.0283		-0.0347	0.0250		0.0197	0.0263	
School size	-0.0010	0.0014		0.0011	0.0012		0.0002	0.0013	
Percent free/reduced-price lunch	0.0038	0.0193		0.0137	0.0169		-0.0257	0.0165	
Percent non-English-speaking	0.0356	0.0345		0.0112	0.0281		0.0135	0.0310	
School has CSR coordinator	-0.0288	0.0150	*	-0.0074	0.0130		-0.0258	0.0128	**
Principal recently joined school	-0.0050	0.0194		0.0014	0.0168		-0.0258	0.0171	
School has middle grades	0.0135	0.0161		-0.0142	0.0143		-0.0111	0.0136	
AYP status 2002–2003	-0.0059	0.0159		0.0052	0.0138		-0.0045	0.0137	
3–5 years of implementation	-0.0116	0.0232		0.0003	0.0203		-0.0207	0.0190	
5 or more years of implementation	-0.0109	0.0272		-0.0248	0.0226		-0.0047	0.0247	
Teacher community: common goals	-0.0227	0.0189		-0.0156	0.0160		-0.0223	0.0157	
Change in teacher community	-0.0009	0.0123		0.0216	0.0106	*	-0.0113	0.0109	
Principal's instructional leadership	0.0075	0.0163		-0.0145	0.0138		-0.0019	0.0133	
Change in principal's instructional leadership	0.0132	0.0117		0.0115	0.0104		0.0039	0.0097	
Usefulness of developer's assistance	0.0178	0.0171		-0.0206	0.0147		-0.0148	0.0138	
Change in usefulness of developer's assistance	0.0008	0.0116		0.0403	0.0099	***	-0.0042	0.0095	
Teacher recently joined school	0.0066	0.0152		0.0044	0.0131		-0.0165	0.0129	
English teacher	-0.0083	0.0139		0.0022	0.0119		-0.0402	0.0148	***
Male teacher	0.0130	0.0186		-0.0050	0.0159		-0.0120	0.0156	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Table 13. Regression Results for Change in Implementation From 2002 to 2004 (continued)

	Time Scheduled for Instruction			Engagement in Informal PD			Influence of Assessments		
Intercept	0.0233	0.0422		0.0482	0.0647		0.0717	0.0479	
CSR Model A key	-0.0297	0.0174	*	-0.0402	0.0274		-0.0506	0.0198	**
CSR Model C key	-0.1023	0.0195	***	0.0731	0.0298	**	-0.0185	0.0213	
CSR Model D/CSR Model G combined key	0.0195	0.0221		0.0398	0.0632		-0.0055	0.0439	
CSR Model E key	-0.0093	0.0245		0.0697	0.0380	*	-0.0757	0.0273	***
CSR Model B key	0.0030	0.0200		0.0104	0.0315		-0.0033	0.0228	
CSR Model F	0.0198	0.0161		0.0249	0.0252		0.0386	0.0184	**
CSR Model A	0.0019	0.0201		0.0346	0.0313		-0.0196	0.0226	
CSR Model C	-0.0728	0.0277	***	-0.0187	0.0418		0.0289	0.0303	
CSR Model D/CSR Model G combined	-0.0143	0.0294		-0.0434	0.0769		0.0304	0.0544	
CSR Model E	0.0431	0.0342		-0.1055	0.0537	*	-0.0105	0.0384	
CSR Model B	-0.0111	0.0324		-0.0092	0.0502		-0.0074	0.0362	
Medium implementation 2002	-0.0128	0.0131		-0.0301	0.0226		-0.0295	0.0163	*
High implementation 2002	-0.0090	0.0164		-0.0451	0.0272	*	-0.0494	0.0195	**
School size	0.0008	0.0008		0.0002	0.0013		-0.0015	0.0009	
Percent free/reduced-price lunch	-0.0076	0.0112		-0.0061	0.0177		-0.0109	0.0127	
Percent non-English-speaking	-0.0242	0.0200		-0.0277	0.0309		0.0216	0.0221	
School has CSR coordinator	0.0020	0.0086		-0.0168	0.0137		-0.0077	0.0099	
Principal recently joined school	-0.0064	0.0111		-0.0078	0.0180		-0.0182	0.0130	
School has middle grades	0.0041	0.0094		0.0035	0.0151		0.0236	0.0108	**
AYP status 2002–2003	-0.0027	0.0091		0.0180	0.0145		-0.0051	0.0105	
3–5 years of implementation	0.0001	0.0134		-0.0050	0.0222		-0.0189	0.0161	
5 or more years of Implementation	-0.0141	0.0158		0.0064	0.0246		-0.0275	0.0178	
Teacher community: common goals	0.0012	0.0108		-0.0196	0.0165		0.0141	0.0124	
Change in teacher community	0.0089	0.0071		0.0054	0.0107		-0.0140	0.0079	*
Principal's instructional leadership	0.0006	0.0094		0.0030	0.0142		-0.0188	0.0105	*
Change in principal's instructional leadership	0.0023	0.0068		0.0536	0.0102	***	0.0143	0.0076	*
Usefulness of developer's assistance	0.0019	0.0099		0.0237	0.0150		0.0030	0.0110	
Change in usefulness of developer's assistance	-0.0011	0.0066		0.0061	0.0099		0.0032	0.0075	
Teacher recently joined school	0.0095	0.0088		-0.0044	0.0133		0.0109	0.0099	
English teacher	-0.0101	0.0080		0.0267	0.0121	**	-0.0043	0.0090	
Male teacher	-0.0074	0.0107		-0.0091	0.0160		-0.0037	0.0121	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

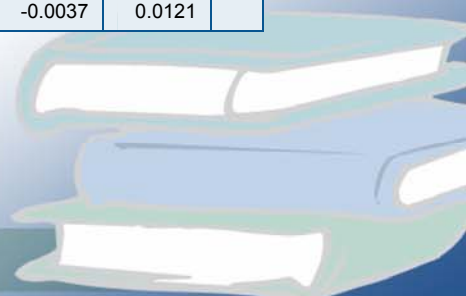


Table 13. Regression Results for Change in Implementation From 2002 to 2004 (continued)

	Use of Assessments		Pedagogy			Emphasis and Type of PD			
Intercept	-0.0054	0.0353		0.0292	0.0397		-0.1618	0.0671	**
CSR Model A key	-0.0213	0.0145		-0.0223	0.0166		-0.0159	0.0286	
CSR Model C key	-0.0101	0.0163		-0.0025	0.0170		-0.0484	0.0320	
CSR Model D/CSR Model G combined key	0.0154	0.0185		-0.0198	0.0340		-0.0764	0.0368	*
CSR Model E key	0.0061	0.0211		0.0036	0.0208		-0.0251	0.0409	
CSR Model B key	-0.0073	0.0169		-0.0176	0.0178		0.0000	0.0340	
CSR Model F	-0.0063	0.0134		-0.0037	0.0170		0.0200	0.0264	
CSR Model A	-0.0026	0.0168		-0.0308	0.0167	*	0.0225	0.0331	
CSR Model C	0.0340	0.0229		-0.0308	0.0214		0.0886	0.0450	*
CSR Model D/CSR Model G combined	-0.0189	0.0245		0.0053	0.0395		0.0299	0.0484	
CSR Model E	0.0008	0.0295		0.0071	0.0281		0.0459	0.0575	
CSR Model B	0.0006	0.0271		-0.0192	0.0261		-0.0125	0.0535	
Medium implementation 2002	-0.0017	0.0110		0.0001	0.0124		-0.0346	0.0217	
High implementation 2002	0.0002	0.0137		-0.0068	0.0162		-0.0688	0.0270	*
School size	-0.0004	0.0007		0.0002	0.0009		0.0017	0.0014	
Percent free/reduced-price lunch	0.0039	0.0094		0.0046	0.0106		0.0207	0.0186	
Percent non-English-speaking	0.0191	0.0167		-0.0218	0.0199		-0.0568	0.0330	*
School has CSR coordinator	-0.0001	0.0073		0.0056	0.0082		-0.0002	0.0143	
Principal recently joined school	-0.0071	0.0093		-0.0118	0.0105		-0.0036	0.0184	
School has middle grades	0.0077	0.0078		0.0023	0.0088		0.0058	0.0155	
AYP status 2002–2003	0.0039	0.0077		-0.0029	0.0086		0.0091	0.0151	
3–5 years of implementation	-0.0134	0.0112		-0.0060	0.0128		-0.0484	0.0221	*
5 or more years of implementation	-0.0017	0.0131		-0.0018	0.0160		-0.0175	0.0258	
Teacher community: common goals	-0.0134	0.0091		0.0152	0.0098		-0.0131	0.0171	
Change in teacher community	-0.0009	0.0059		-0.0083	0.0066		-0.0001	0.0111	
Principal's instructional leadership	-0.0095	0.0078		-0.0172	0.0084	*	0.0527	0.0149	***
Change in principal's instructional leadership	-0.0012	0.0057		0.0003	0.0059		0.0374	0.0107	***
Usefulness of developer's assistance	0.0229	0.0083	***	0.0008	0.0086		0.0249	0.0158	
Change in usefulness of developer's assistance	-0.0012	0.0055		0.0033	0.0059		0.0176	0.0104	*
Teacher recently joined school	0.0032	0.0074		-0.0057	0.0081		-0.0090	0.0139	
English teacher	0.0107	0.0067		0.0082	0.0086		0.0058	0.0126	
Male teacher	0.0136	0.0090		0.0035	0.0095		0.0050	0.0169	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Discussion

The goal of this paper was to explore how implementation of CSR models varies and what predicts the level of and change in CSR model implementation. We were particularly interested in exploring four specific research questions:

1. Are schools that implement CSR models engaged in different sets of practices and activities regarding components of CSR models than schools that are not implementing CSR models?
2. Does the level of implementation vary among CSR models?
3. Does the level of implementation vary by component and by how comprehensively CSR models are implemented?
4. What factors other than having a specific CSR model predict the level and change in implementation?

To answer question one, we needed to be able to include in our analysis comparison schools that did not implement any CSR programs, because we do not have data for schools implementing a CSR model prior to the adoption of the CSR model. Our conceptualization of implementation—as the set of practices that a fully implementing CSR model school should be engaged in—made this possible. CSR program developers provided us with implementation keys (principal surveys and teacher surveys filled out as if a school/teacher were fully implementing their respective CSR model). We used the developer’s survey answers to create implementation indices both for schools actually implementing a CSR model and for their matched comparison schools.⁸ Therefore, we were able to compare the level of implementation between schools implementing CSR models and their respective comparison schools.

Our four major findings are summarized below.

- ◆ Finding 1: CSR schools do not systematically have higher levels of implementation than their matched comparison schools.
 - According to our results, CSR models are engaged in different practices than their paired comparison schools, but the results are not consistent. Our descriptive results showed only a few statistically significant results. Cross-sectional HLM analysis for years 2002 and 2004 revealed additional differences, but the results were not consistent. Sometimes, CSR model schools had a higher fidelity of implementation than their comparison schools, while occasionally comparison schools implemented more faithfully. The results were also unstable over time: what was significant in 2002 did not remain so in 2004.
 - Some of the differences that did exist conform to our pre-existing knowledge of the CSR models. CSR Model F schools had a higher level of implementation in Use of Assessments, whereas CSR Model C schools had higher fidelity of implementation in Use of Technology and Pedagogy (in 2002). CSR Model A schools did better in implementing Pedagogy than comparison schools in 2002, but this relationship was reversed in 2004.
 - The small number of significant findings and their instability could be due to several factors. First, all schools, regardless of whether they are implementing a CSR model or not, may be

⁸ Matching was done by propensity scores. See Appendix B for a detailed description.



engaged in a core set of practices that do not vary tremendously from school to school. This lack of differences may be because many practices initiated and promoted by different CSR models have been mainstreamed over time. Alternatively, it is possible that, although all schools were engaged in the same practices, these practices should be conducted in CSR schools with a qualitative difference. Our implementation indices based on survey data do capture how often or what level of emphasis should be given to a certain practice, but the data do not include more nuanced information about how a certain practice/activity should be conducted. Second, implementing a CSR model is not likely to be a linear process in which all components of CSR are simultaneously highly implemented. It is likely that schools engaged in selective implementation and concentrated on different aspects of CSR models. Thus, only some components of a CSR model may be highly implemented at a specific time point. Over time, however, a trend of increased levels of implementation should be detected for most components of CSR models. We did detect an increase in the level of implementation over time (between 2002 and 2004), but this trend took place in both CSR model schools and comparison schools. This common trend explains why we did not detect more differences in 2004 between CSR schools and their matched comparison schools.

- ◆ Finding 2: CSR model keys are significant predictors for the level and change of implementation.
 - Although we found few significant differences between CSR model schools and their matched comparison schools, large differences exist between CSR model keys concerning the implementation of different components of CSR models. In other words, the level of implementation is clearly related to the bundle of activities each CSR model recommends or requires as part of its implementation. This effect became especially clear, because the results were consistent over time (between 2002 and 2004). That is, certain CSR models tended to have lower levels of implementation concerning specific components of that CSR model.
 - In our analysis, the CSR Model F and CSR Model E model keys consistently had medium to high levels of implementation. The CSR Model B model key was related to a lower level of implementation regarding Shared Decision Making, Use of Technology, and Inclusion, whereas the CSR Model A model key had particularly low levels of implementation regarding Inclusion and Pedagogy. CSR Model C had the lowest levels of implementation in Student Grouping and Time Scheduled for Instruction. Clearly, the model developer's keys for implementation made a difference regarding the level of implementation. A closer look at the components in which particular CSR programs had lower levels of implementation reveals that the CSR model implementation keys had different levels of difficulty. In other words, some CSR developers demanded practices and activities that are not common in most schools, requiring a *change* in ongoing school operations. Thus, the level of that implementation tended to be lower.
 - This finding raises an interesting question regarding the effectiveness of CSR models promoting positive change in student achievement. Assuming that profound change in the school operations, from governance to instruction, is required to have a positive effect on student achievement, it is not sufficient that a CSR model is highly implemented if the model does not promote real change. A study of which CSR model key promotes positive change in student achievement is beyond the scope of this paper, but the issue demands further inquiry.



- ◆ Finding 3: CSR models are not comprehensively implemented.
 - CSR is supposed to be comprehensive, addressing several dimensions of school operations (governance, PD, assessments, and instruction, among others). Our results showed that about one fifth of the schools in the study implemented CSR models comprehensively in 2002, while only about one tenth did so in 2004. However, schools implementing a CSR model were more likely to have high a comprehensiveness level than comparison schools, in both 2002 and 2004. The low level of comprehensiveness may be due to selective implementation: schools may decide to implement only particular components of the CSR program, or schools may decide to implement different components sequentially, one after another. Thus, implementation of a CSR model is a dynamic process in which the level of implementation of different model components varies at different times, but the comprehensiveness of CSR implementations should increase over time. In our sample, the comprehensiveness decreased over time, but this finding is likely caused by differentiated attrition (many schools with high levels of comprehensiveness in 2002 were missing in the 2004 sample).
- ◆ Finding 4: The level and change in implementation is more consistently predicted by variables related to agency (such as Principal's Instructional Leadership and Developer's Assistance) than to contextual variables.
 - Our HLM analysis illustrates what factors predict the level and change in implementation. As discussed above, CSR model keys significantly predict the level of implementation for all the implementation indices. Our results also showed that Principal's Instructional Leadership, Developer's Assistance, and Teachers' Professional Community were consistently and positively related to the level of implementation across many implementation indices. Similarly, being an English teacher (instead of being a mathematics teacher) was positively related to implementation across the implementation indices that were analyzed. School-level characteristics (such as Percentage of Students Receiving Free/Reduced-Price Lunch, Percentage of Non-English-Speaking Students, and School Size), although significantly related to some of the implementation indices, did not consistently predict level of implementation. These results illustrate the importance of agents (teachers, principals, model developers) in the implementation process instead of the context: CSR models can be successfully implemented in different environments if the relevant actors are engaged in the process of implementation.
 - Contextual variables were not consistently associated with the change in implementation, except for one variable: positive change in Principal's Instructional Leadership increases the level of implementation. This finding again reinforces the importance of the principal's leadership in CSR model implementation. Measures for Teacher Community and Developer's Assistance did not predict the change in implementation systematically across different implementation indices.



References

- Aladjem, D. K. (2003). *Measuring implementation across multiple interventions*. Paper presented at the Association for Public Policy Analysis and Management Annual Research Conference, Washington, DC.
- Barnes, C., Camburn, E., Kim, J., & Rowan, B. (2004). *School leadership and instructional improvement in CSR schools*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Berends, M. (2000). Teacher-reported effects of New American Schools design: Exploring relationships to teacher background and school context. *Educational Evaluation and Policy Analysis, 22*(1), 65–82.
- Berends, M., Bodilly, S., & Kirby, S. N. (2002). Facing the challenges of whole-school reform: New American Schools after a decade. Santa Monica, CA: RAND.
- Berends, M., Kirby, S. N., Naftel, S., & McKelvey, C. (2001). Implementation and performance in New American Schools: Three years into scale-up. Santa Monica, CA: RAND Corporation.
- Bifulco, R., Duncombe, W., & Yinger, J. (2005). Does whole-school reform boost student performance: The case of New York City. *Journal of Policy Analysis and Management 24*(1), 47–72.
- Bloom, H. S., Rock, J., Ham, S., Melton, L., & O'Brien, J. (2001). *Evaluating the Accelerated Schools approach*. New York: MDRC.
- Bodilly, S. J., Keltner, B., Purnell, S., Reichardt, R., & Schuyler, G. (1998). Lessons from New American Schools' scale-up phase: Prospects for bringing designs to multiple schools. Santa Monica: RAND.
- Cook, T. D., Habib, F., Phillips, M., Settersten, R. A., Shagle, S. C., & Degirmencioglu, S. M. (1999). Comer's School Development Program in Prince George's County: A theory-based evaluation. *American Educational Research Journal 36*(3), 543–597.
- Cooper, R., Slavin, R. E., & Madden, N. A. (1998). Success For All: Improving the quality of implementation of whole-school change through the use of a national reform network. *Education and Urban Society, 33*(3), 385–408.
- Datnow, A., Borman, G., Stringfield, S., Rachuba, L., & Castellano, M. (2003). Comprehensive school reform in culturally and linguistically diverse contexts: Implementation and outcomes from a four-year study. *Educational Evaluation and Policy Analysis, 25*(2), 25–54.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 72*(3), 433–479.



References (continued)

- Engelmann, S. E., & Engelmann, K. E. (2004). Impediments to scaling up effective comprehensive school reform models. In T. K. Glennan, Jr., S. Bodilly, J. Galegher, & K. A. Kerr (Eds.), *Expanding the reach of education reforms: Perspectives from leaders in the scale-up of educational interventions* (pp. 107–133). Santa Monica, CA: RAND.
- Hatch, T. (1998). The differences in theory that matter in the practice of school improvement. *American Educational Research Journal*, 35(1), 3–31.
- Kirby, S. N., Berends, M., Naftel, S., & Sloan, J. S. (2001). Schools adopting Comprehensive School Reform Demonstration models: Early findings on implementation. Santa Monica, CA: RAND.
- Slavin, R., & Madden, N. A. (1999). *Success for All/Roots and Wings: Summary of research on achievement outcomes*. Baltimore, MD: Center for Research on the Education for Students Placed at Risk.
- Smith, J., Maxwell, S., Lowther, D., Hacker, D., Bol, L., & Nunery, J. (1997). Activities in schools and programs experiencing the most, and least, early implementation successes. *School Effectiveness and School Improvement*, 8(1), 125–150.
- Smith, L., Ross, S., McNelis, M., Squires, M., et al. (1998). The Memphis Restructuring Initiative: Analysis of activities and outcomes that impact implementation. *Education and Urban Society*, 30(3), 296–325.
- Stringfield, S., & Datnow, A. (1998). Introduction: Scaling up school restructuring designs in urban schools. *Education and Urban Society*, 30(3), 269–276.
- Supovitz, J. A., & May, H. (2003). *The relationship between teacher implementation of America's Choice and student learning in Plainfield, New Jersey*. Philadelphia: Consortium for Policy Research in Education.
- Yonezawa, S., & Datnow, A. (1999). Supporting multiple reform designs in a culturally and linguistically diverse school district. *Journal of Education for Students Placed at Risk*, 4(1), 101–125.



Appendix A: Implementation Indices

This section describes how our implementation measure actually works. More specifically, this section focuses on how we calculate implementation for each school in our sample.

The implementation-as-fidelity idea is predicated on the idea that we can measure implementation across our models by comparing the CSR program implementation in each school with the concept of the fully implemented program as described by the developers' responses to the surveys. We want to compare the bundle of activities or the configuration of CSR components as they exist in the schools in our sample with what the developers would expect a school to be doing with full implementation of the program.

Conceptually, that is, we have the schools in our sample doing one set of things, and we want to compare those sets with the corresponding "right" set to determine the level of alignment/similarity. The more aligned or closer a given school is to its "ideal," the more fully implemented it is. So implementation (I) is a function of the developers' normative vision of school organization and activities (D_k for each k developer) and the empirical practice of schools in our sample (S_i for each school i).

$$I = f(D_k, S_i) \quad (1)$$

Because we have survey data from mathematics teachers, English/language arts teachers, and principals, we can expand Equation 1 to Equation 2, substituting M , E , and P for S , representing mathematics teachers, English/language arts teachers, and principals, respectively.

$$I = f(D_k, M_i, E_i, P) \quad (2)$$

Although many items differ with respect to each content area, we have constructed our scales such that the mathematics and English/language arts teacher surveys can be considered the same, and we combine M and E from Equation 2 into T . (For example, a scale representing instructional style measures the extent to which teachers employ more constructivist or didactic pedagogies. Although the items making up the scale differ across the mathematics and English/language arts surveys, they differ only in the content area-specific examples used.) Equation 3 now represents how we conceptualize implementation as a function of the developers' normative vision and the empirical reality experienced by teachers and principals.

$$I = f(D_k, T_i, P) \quad (3)$$

Operationally, however, we need to partition implementation into teacher and principal segments, as shown in Equations 4 and 5.

$$I_T = f(D_k, T_i) \quad (4)$$

$$I_P = f(D_k, P) \quad (5)$$



Equations 4 and 5 lead to the question of how to operationalize f . To operationalize f , we must consider the available data. Our survey data (including the developer responses)—both the scales and individual items—can be abstracted as follows:

With only one principal per school, the principal data can be represented by \mathbf{p} , an n -dimensional row vector, where n = the number of scales and items to be included.

The developer norms for teachers can be defined as \mathbf{d}_k^t , an n -dimensional row vector for each k program developer, where n = the number of scales and items to be included.

The developer norms for principals constitute \mathbf{d}_k^p , an n -dimensional row vector for each k program developer, where n = the number of scales and items to be included.

Calculating Implementation: Euclidean Distance

If implementation is a function of the developers' normative vision and the empirical experiences of principals and teachers, one of the simplest ways to conceptualize the difference between the "actual" and "ideal" levels of implementation is to use Euclidean distance measure. For this paper, we used squared Euclidean distance. Euclidean distance gives a measure of dissimilarity of two vectors in multidimensional space.

$$\text{Level of implementation} = \sum (X_{\text{Ideal}_i} - Y_{\text{Actual}_i})^2 / \text{number of survey items used in index}$$

All variables used in the calculation of the implementation index were standardized, because Euclidean distance measures are sensitive to the scales of included variables. Similarly, to make implementation indices for different CSR components more comparable, we divided the distance measure by the number of variables included in each implementation index.

Although this approach gives a measure of dissimilarity of two vectors, interpretation of the results is not straightforward. It is unclear how much worse the implementation is in a school with a distance 5.7 is than in a school having a value of 4.2. To overcome this problem, we used the developers' survey answers to construct a hypothetical school with the worst possible level of implementation for each implementation index for each CSR program. We used these hypothetical schools to calculate percentage of implementation variable for each CSR component:

$$\text{Percentage of Implementation} = (1 - [\text{dissimilarity of Ideal and Actual for school } X_i / \text{For Component } i \text{ distance between Ideal and Hypothetical School } X_{\text{max}}])$$

We were able to calculate all implementation indices for all CSR programs and indices for which we received a developer's key. However, it should be noticed that one developer did not return a survey for English teachers; therefore, some teachers are missing both Curriculum and Pedagogy indices. In addition, another developer did not provide a key for three indices (Technology, Inclusion, and Parental Involvement), while a third developer did not answer a question required for creating of Inclusion, Parental Involvement, Curriculum, Engagement in Informal PD, Influence of Assessments, and Pedagogy. To avoid losing information if a CSR developer or principals and teachers did not answer



either teacher or principal survey questions, we combined principal and teacher measures when they measured similar concepts. Such combination of measures was done for Parent/Community Involvement, Inclusion, and Shared Decision Making.



Appendix B: Propensity Scores

To compare the level of CSR program implementation between schools implementing CSR and control schools, we must find a way to calculate implementation indices for control schools included in the study. We have developer’s surveys and surveys from participating principals and teachers, but to calculate implementation indices for control schools, we have to first decide with which CSR program each control school should be matched.

We decided to use a propensity score approach to match control schools to a specific CSR program to calculate implementation indices for control schools. The basic idea of the propensity score approach is quite simple: we calculated the likelihood for each school to be a CSR school (their propensity of being a CSR school) based on school characteristics. Then we matched, within a school district, a control school to a CSR school that had a similar propensity of being a CSR school. Thus, if a control school had a similar propensity of being a CSR school as a school implementing CSR program “C,” the control school would be “assigned” CSR program C in order to calculate its level of implementation.

Propensity scores were calculated for control schools and for schools implementing one of the seven CSR programs for which we have developers’ surveys. As a result, propensity scores were calculated for 503 schools; the actual number of schools included in the implementation analysis varies, depending on whether principals and teachers returned their survey instruments.

We used logistic regression to calculate propensity scores for all 3 years of our study (2002–2004). We used scores from 2003 and 2004 only if scores for year 2002 were missing. The logistic regression model included the following variables: dummy variables for school districts (excluding two districts that did not have control schools), school enrollment, percentage of children receiving free/reduced-price lunch, percent of African American students, percentage of non-English-speaking students, school’s IFI status, principal’s gender, principal’s experience (years in this school), principal’s race, a dummy variable for whether a school is a regular public school, a combined variable for school-level mathematics achievement (from years 1999–2002), interaction terms between mathematics achievement and CSR concentration (percentage of CSR schools in district), interaction between challenging environment and whether the district requires schools to adopt CSR programs, interaction between mathematics achievement and whether the district requires schools to adopt CSR programs, emphasis on noninstructional goals, and use of state and district assessment data in decision making.

For each year of the study (from year 1 to year 3), the percentages of concordant pairs were 73.5%, 71.0%, and 75.0%, respectively.

Control schools were matched to CSR schools within each school district. Two decision rules were used in the matching process: (a) schools could be matched only if they had less than a 10-point difference (most of the time, the difference is much smaller), and (b) because CSR program “A” is overrepresented in the sample, we would not match a control school to CSR program A if another CSR program was available. The resulting distribution of schools in seven different CSR programs is shown in Table B1.



Table B1. Number of CSR Schools and Control Schools Matched to Specific CSR Programs

CSR model	Year 1		Year 3	
	CSR schools	Comparison schools	CSR schools	Comparison schools
CSR Model A	33	52	29	48
CSR Model B	13	19	12	16
CSR Model C	27	33	22	28
CSR Model F	85	93	80	87
CSR Model D	9	18	8	15
CSR Model E	12	11	9	10
CSR Model G	6	4	6	2
Total	185	230	166	206



Appendix C: Survey Questions Used to Create Implementation Indices

Table C1. Survey Questions Used to Create Implementation Indices

Governance			
Shared Decision Making	Principal	(Q1) Please indicate the relative emphasis placed on each of these goals/strategies within your school this year. <i>h. Sharing decision-making authority among faculty, staff, and administrators</i>	<u>Response scale:</u> 1 = No emphasis/ not needed 2 = Minor emphasis 3 = Moderate emphasis 4 = Major emphasis
		(Q6) How much influence do the district, school committee, principal, and/or individual teachers have on the following decisions or activities? <i>a. Hiring new teachers</i> <i>b. Selecting professional development activities for teachers</i> <i>c. Planning school budgets</i> <i>d. Developing a school improvement plan or action plan</i> <i>e. Selecting textbooks</i>	<u>Response scale:</u> 1 = Mostly a district decision 2 = Mostly a school decision 3 = Evenly shared district-school responsibility 4 = Mostly a classroom teacher responsibility
	Teacher	(Q4) Please indicate the emphasis placed on each of these goals/strategies within your school this year. <i>h. Sharing decision-making authority among staff and administrators</i>	<u>Response scale:</u> 1 = No emphasis/ not needed 2 = Minor emphasis 3 = Moderate emphasis 4 = Major emphasis



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Governance (continued)			
<i>Shared Decision Making</i>	Teacher	(Q5) How much influence do the district, school committee, principal, and individual teachers have on the following decisions? <i>a. Selecting instructional materials</i> <i>b. Selecting topics and skills to be taught</i> <i>c. Selecting teaching techniques</i> <i>d. Creating student ability groups for instruction in each classroom</i> <i>e. Allocating instructional time for each academic subject</i>	<u>Response scale:</u> 1 = Mostly a district decision 2 = Mostly a school decision 3 = Evenly shared district-school responsibility 4 = Mostly a classroom teacher responsibility
		(Q23) Since September 2000, how frequently did you engage in each of the following activities for English/language arts or mathematics? <i>e. Participating in a learning community (teacher collaboratives, networks, or study groups)</i> <i>h. Participating in a committee or task force focused on curriculum and instruction</i>	<u>Response scale:</u> 1 = Never 2 = A few times a year 3 = Once or twice a month 4 = Once or twice a week 5 = Almost Daily
<i>School Organization</i>	Principal	(Q3) Does your school have a team, committee, council, or other such group to make decisions or recommendations for the school? 1. <i>Yes, this group makes decisions</i> 2. <i>Yes, this group makes recommendations</i> (Q4) Who is involved in this group? 3. <i>Parents</i> 4. <i>Instructional specialists (e.g., bilingual, special education)</i> 5. <i>Students</i> (Q5a) Does your school have a designated person who acts as a coordinator, facilitator, or coach for school reform?	<u>Response options:</u> 1 = Yes 0 = No



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Technology			
<i>Use of Technology in Classrooms</i>	Teacher	(Q8) During instruction, how frequently do your students in your target English/ language arts or mathematics class use computers to do the following? <i>a. Use computer applications such as word processing, spreadsheets, etc.</i> <i>b. Practice drills</i> <i>c. Research using the Internet and/or CD-ROM</i> <i>d. Take assessments</i>	<u>Response scale:</u> 1 = Never 2 = Once or twice a semester 3 = Once or twice a month 4 = Once or twice a week 5 = Almost Daily
Parent/Community Involvement			
<i>Parent/Community Involvement</i>	Principal	(Q1) Please indicate the relative emphasis placed on each of these goals/strategies within your school this year. <i>d. Ensuring open communication between the school and parents</i> <i>e. Promoting community involvement in school activities</i> <i>m. Attending to student health issues</i>	<u>Response scale:</u> 1 = No emphasis/ not needed 2 = Minor emphasis 3 = Moderate emphasis 4 = Major emphasis
	Teacher	(Q29) If your school offers any of the following activities or events, please indicate how many of <i>your students'</i> parents participate. <i>a. Teacher-parent conferences</i> <i>b. Sign daily activity sheet for completion of homework</i> <i>c. Home visits from teacher or other staff</i> <i>d. Science fairs, math nights, or other academic activity for students and parents</i> <i>e. Develop written agreements between the school and parents that describe what each will do to help students succeed.</i> <i>f. E-mail/phone call communication</i>	<u>Response scale:</u> 1 = Not available in our school 2 = Just a few parents 3 = About a quarter of parents 4 = About half of parents 5 = Most parents

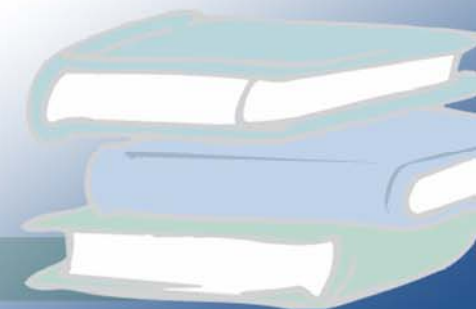


Table C1. Survey Questions Used to Create Implementation Indices (continued)

Parent/Community Involvement (continued)			
<i>Parent/Community Involvement</i>	Teacher	<p>(Q30) How often do you communicate information to the parents of your students in the following ways?</p> <p><i>a. Inform parents about learning objectives in core academic subjects</i></p> <p><i>b. Contact parents when their child is encountering academic problems</i></p> <p><i>c. Assign homework or provide activities that require children to interact with parents</i></p>	<p><u>Response scale:</u></p> <p>1 = Never</p> <p>2 = A few times a year</p> <p>3 = Once or twice a month</p> <p>4 = Once or twice a week</p> <p>5 = Almost Daily</p>
Professional Development			
<i>Emphasis of Professional Development</i>	Teacher	<p>(Q27) How much emphasis did your professional development activities place on the following topics?</p> <p><i>a. State or district content standards in English/language arts</i></p> <p><i>b. Curriculum associated with a specific English/language arts program</i></p> <p><i>c. Student assessment techniques in English/language arts</i></p> <p><i>d. Instructional strategies for non-native English-speaking students</i></p> <p><i>e. Instructional strategies for special education students</i></p> <p><i>f. Instructional strategies for low-achieving students</i></p> <p><i>g. Using student work to think about changing instruction or curricula</i></p> <p><i>h. Using drills, memorization or other skills-based activities</i></p> <p><i>i. Using long-term projects or reports</i></p> <p><i>j. Relating lessons to the real world</i></p> <p><i>k. Using cooperative learning techniques</i></p> <p><i>l. Technology</i></p> <p><i>m. Classroom management</i></p> <p><i>n. School management or governance</i></p> <p><i>o. Data-based decision making</i></p>	<p><u>Response scale:</u></p> <p>1 = No emphasis</p> <p>2 = Minor emphasis</p> <p>3 = Moderate emphasis</p> <p>4 = Major emphasis</p>



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Professional Development (continued)			
<i>Emphasis of Professional Development</i>	Teacher	<p>(Q26) What changes have you made in the following areas of your teaching practice as a result of your professional development activities in English/language arts or mathematics since September 2000 [2002 in the 2003–2004 survey]?</p> <p><i>d. The types or mix of assessments I use to evaluate students</i></p>	<p><u>Response scale:</u></p> <p>1 = No change 2 = Minor change 3 = Moderate change 4 = Major change</p>
<i>Type of Professional Development</i>	Teacher	<p>(Q24) Teachers may participate in professional development activities alone or with groups of teachers from their school. Since September 2000 [2002 in the 2003–2004 survey], how often did you participate in professional development activities in English/language arts or mathematics in the following ways?</p> <p><i>a. I participated in professional development with most or all of the teachers in my school</i></p> <p><i>b. I participated in professional development with most or all of the teachers in my department or grade level</i></p> <p>(Q25) Thinking again about all of your professional development activities in English/language arts or mathematics since September 2000 [2002 in the 2003–2004 survey], how often did the following occur?</p> <p><i>a. Participants observed demonstrations of teaching techniques</i></p> <p><i>b. Participants practiced what they learned and received feedback</i></p> <p><i>c. Participants led group discussions</i></p> <p><i>d. Participants conducted a demonstration of a lesson, unit, or skill</i></p> <p><i>e. Participants developed and practiced using student materials</i></p> <p><i>f. Participants reviewed student work or scored assessments</i></p>	<p><u>Response scale:</u></p> <p>1 = Never 2 = Rarely 3 = Sometimes 4 = Often</p>

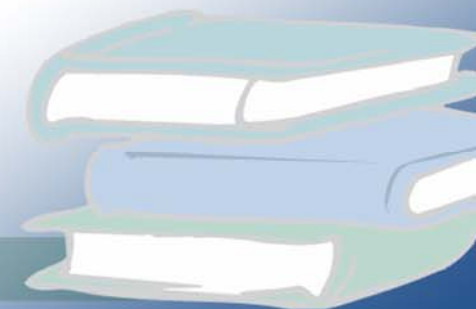


Table C1. Survey Questions Used to Create Implementation Indices (continued)

Professional Development (continued)			
<i>Engagement in Informal Professional Development</i>	Teacher	<p>(Q4) Please indicate the emphasis placed on each of these goals/strategies within your school this year.</p> <p><i>i. Ensuring our teachers attend high-quality professional development activities</i></p> <p>(Q27) How much emphasis did your professional development activities place on the following topics?</p> <p><i>a. State or district content standards in English/language arts</i></p> <p><i>b. Curriculum associated with a specific English/language arts program</i></p> <p><i>c. Student assessment techniques in English/language arts</i></p> <p><i>d. Instructional strategies for non-native English-speaking students</i></p> <p><i>e. Instructional strategies for special education students</i></p> <p><i>f. Instructional strategies for low-achieving students</i></p> <p><i>g. Using student work to think about changing instruction or curricula</i></p> <p><i>h. Using drills, memorization or other skills-based activities</i></p> <p><i>i. Using long-term projects or reports</i></p> <p><i>j. Relating lessons to the real world</i></p> <p><i>k. Using cooperative learning techniques</i></p> <p><i>l. Technology</i></p> <p><i>m. Classroom management</i></p> <p><i>n. School management or governance</i></p> <p><i>o. Data-based decision making</i></p>	<p><u>Response scale:</u></p> <p>1 = No emphasis</p> <p>2 = Minor emphasis</p> <p>3 = Moderate emphasis</p> <p>4 = Major emphasis</p>



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Assessment			
<i>Influence of Assessments</i>	Teacher	<p>(Q19) In your target classroom, what influence does each of the following types of <i>classroom assessments</i> have on a student's final grade?</p> <p>a. <i>Multiple-choice questions on tests</i></p> <p>b. <i>Essays, short-answer questions, or other writing assignments</i></p> <p>c. <i>Portfolio of student work</i></p> <p>d. <i>Group projects and presentations</i></p> <p>e. <i>Individual student demonstrations, exhibitions, or oral presentations</i></p>	<p><u>Response scale:</u></p> <p>1 = No influence</p> <p>2 = Minor Influence</p> <p>3 = Moderate Influence</p> <p>4 = Major Influence</p>
<i>Use of Assessments</i>	Teacher	<p>(Q20) How strongly do you agree or disagree with the following statements about classroom assessments in your target English/language arts class?</p> <p>a. <i>Classroom assessments are continuous and explicitly linked to subject matter taught</i></p> <p>b. <i>Students are provided with a rubric or guidelines that explain how assessment tasks will be evaluated</i></p> <p>c. <i>Assessments which I personally develop are a significant part of my instruction.</i></p> <p>d. <i>Assessments from external sources are a significant part of my instruction.</i></p>	<p><u>Response scale:</u></p> <p>1 = Strongly disagree</p> <p>2 = Disagree</p> <p>3 = Agree</p> <p>4 = Strongly Agree</p>

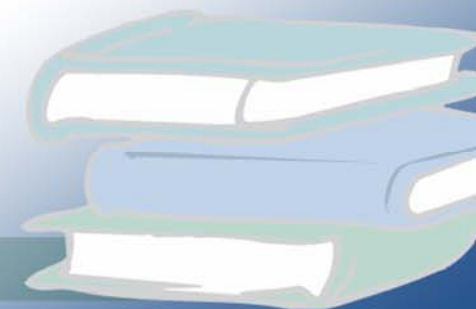


Table C1. Survey Questions Used to Create Implementation Indices (continued)

Organization of Teaching/Classrooms			
<i>Inclusion</i>	Principal	<p>(Q10) I am now going to ask a few questions concerning the grouping in your school of non-native English speakers (sometimes referred to as LEP, ESOL, or ELL) as well as students with disabilities during the school day.</p> <p><i>A1. Are non-native English speakers mainstreamed in general education classes?</i></p> <p><i>A2. Are students with disabilities mainstreamed in general education classes?</i></p> <p><i>B1. Are non-native English speakers grouped in self-contained classes?</i></p> <p><i>B2. Are students with disabilities grouped in self-contained classes?</i></p> <p><i>C1. Do non-native English speakers participate in pull-out instruction for part of the school day?</i></p> <p><i>C2. Do students with disabilities participate in pull-out instruction for part of the school day?</i></p>	<p><u>Response options:</u></p> <p>1 = Yes</p> <p>0 = No</p>



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Organization of Teaching/Classrooms (continued)			
<i>Inclusion</i>	Teacher	<p>(Q16) What statement best characterizes the books, materials, or resources you use with LEP or ESOL students who are a part of your target English/language arts class? (Circle all that apply)</p> <ol style="list-style-type: none"> 1. <i>LEP or ESOL students use the same materials as general education students</i> 2. <i>Books, materials, and resources are written in the native language of the LEP or ESOL students</i> 3. <i>Books, materials, and resources are tailored to the instructional level of the LEP or ESOL students</i> 4. <i>LEP or ESOL specialists work with students</i> <p>(Q18) Which statement best characterizes the books, materials, or resources you use with students with disabilities (who have current IEPs) who are a part of your target English/language arts class? (Circle all that apply)</p> <ol style="list-style-type: none"> 1. <i>Students with disabilities use the same books, materials, and resources (without adaptations) as do general education students</i> 2. <i>Students with disabilities use books, materials, and resources that are adapted for their special needs, but with the same curricular content as general education students</i> 3. <i>Students with disabilities use books, materials, and resources that are adapted for their special needs, with different curricular content</i> 4. <i>Specialists work with students with disabilities</i> 	<p><u>Response options:</u></p> <p>1 = Yes 0 = No</p>
<i>Student Grouping</i>	Teacher	<p>(Q9) When teaching, how often do you use the following approaches to group students for instruction in your target English/language arts or mathematics class?</p> <ol style="list-style-type: none"> a. <i>Similar ability or achievement level</i> b. <i>Mixed ability or achievement level</i> 	<p><u>Response scale:</u></p> <p>1 = Never 2 = Rarely 3 = Sometimes 4 = Often</p>

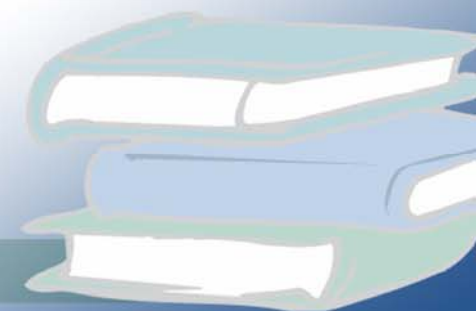


Table C1. Survey Questions Used to Create Implementation Indices (continued)

Organization of Teaching/Classrooms (continued)			
<i>Student Grouping</i>	Teacher	(Q10) How frequently are student grouping patterns reevaluated within your target English/language arts or mathematics class?	<p><u>Response options:</u></p> <p>1 = About once a week 2 = About once a month 3 = Every 2 months 4 = Every semester 5 = Students mostly remain with the same group all year 6 = As necessary, not on a regular basis</p>
<i>Time Scheduled for Teaching</i>	Teacher	<p>How often do children in your class(es) usually work on lessons or projects in the following general topic areas, whether as a whole class, in small groups, or in individualized arrangements?</p> <p>7a. <i>English and language arts</i> 7b. <i>Mathematics</i> 7c. <i>Social studies</i> 7d. <i>Science</i></p>	<p><u>Response scale:</u></p> <p>1 = Never 2 = Less than once a week 3 = 1-2 times a week 4 = 3-4 times a week 5 = Daily</p>
		<p>How much time do children in your class(es) usually work on lessons or projects in the following general topic areas, whether as a whole class, in small groups, or in individualized arrangements?</p> <p>7a_1 <i>English and language arts</i> 7b_1 <i>Mathematics</i> 7c_1 <i>Social studies</i> 7d_1 <i>Science</i></p>	<p><u>Response scale:</u></p> <p>1 = Less than 40 minutes a session 2 = 40-55 minutes a session 3 = 56-70 minutes a session 4 = More than 70 minutes a session</p>



Table C1. Survey Questions Used to Create Implementation Indices (continued)

Instruction				
Curriculum	Teacher	English	<p>(Q11) This year, what emphasis did you place on the following topics in your target English/language arts class?</p> <ul style="list-style-type: none"> a. <i>Word analysis</i> b. <i>Vocabulary development</i> c. <i>Reading comprehension</i> d. <i>Writing development</i> e. <i>Textual features</i> f. <i>Literature – fiction</i> g. <i>Literature – nonfiction</i> h. <i>Information/study skills</i> i. <i>Communication skills</i> 	<p><u>Response scale:</u></p> <ul style="list-style-type: none"> 1 = No emphasis 2 = Minor emphasis 3 = Moderate emphasis 4 = Major emphasis
		Math	<p>(Q11) This year, how much emphasis did you place on the following topics in your mathematics instruction?</p> <ul style="list-style-type: none"> a. <i>Number sense and numeration</i> b. <i>Whole numbers</i> c. <i>Fractions</i> d. <i>Decimals</i> e. <i>Percent</i> f. <i>Ratio/proportions</i> g. <i>Measurement</i> h. <i>Pre-algebra</i> i. <i>Algebra</i> j. <i>Geometry</i> 	<p><u>Response scale:</u></p> <ul style="list-style-type: none"> 1 = No emphasis 2 = Minor emphasis 3 = Moderate emphasis 4 = Major emphasis

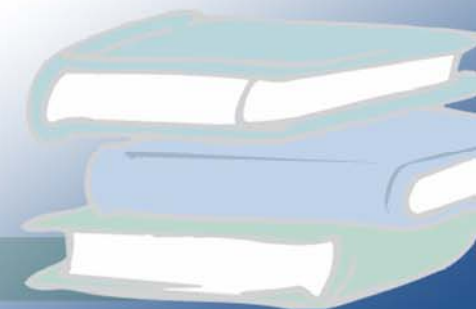


Table C1. Survey Questions Used to Create Implementation Indices (continued)

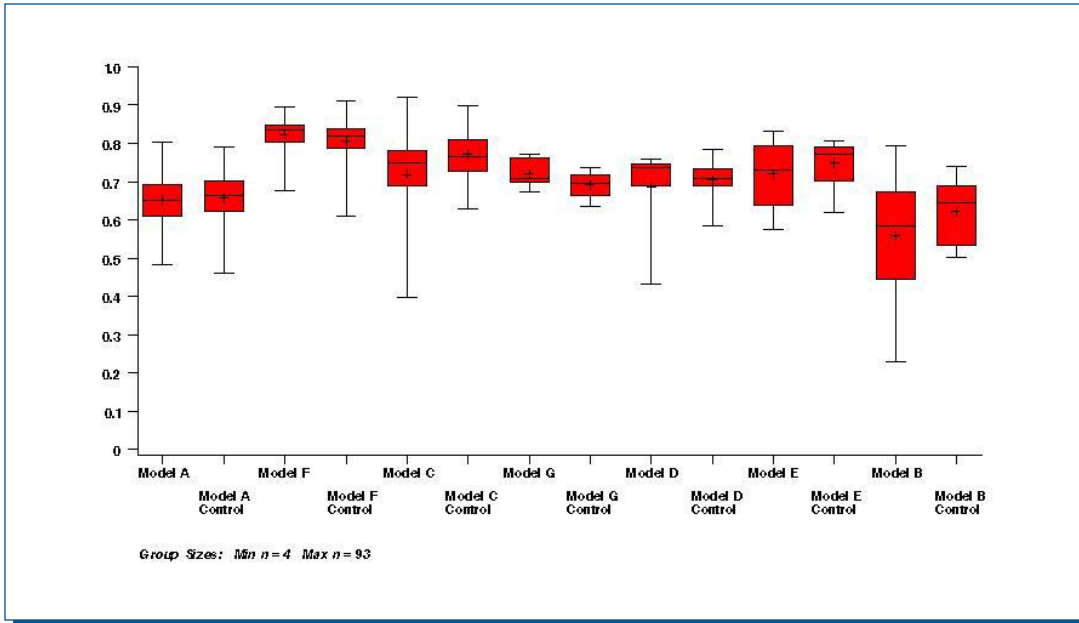
Instruction (continued)				
Pedagogy	Teacher	English	<p>(Q12) This year, how often did students in your target English/language arts class do the following?</p> <ul style="list-style-type: none"> a. Listen to me give formal presentations of definitions or concepts. b. Write brief answers to questions about something they have read c. Work on a written product or report for several days d. Make predictions about what they are reading as they are reading it e. Explain, support, or justify their understanding of what they have read f. Work in a reading workbook or on a worksheet g. Make inferences or use other analytic reading strategies to explore what they have read h. Take a written quiz or test about what they have read 	<p><u>Response scale:</u></p> <ul style="list-style-type: none"> 1 = Never 2 = Once or twice a semester 3 = Once or twice a month 4 = Once or twice a week 5 = Almost every day
		Math	<p>(Q12) This year, how often did students in your target mathematics class do the following?</p> <ul style="list-style-type: none"> a. Listen to me present the definition of a term or the steps of a procedure b. Perform tasks requiring methods or ideas already introduced to students c. Assess a problem and choose a method to use from those already introduced to students d. Perform tasks requiring methods or ideas not already introduced to students e. Explain an answer or solution method for a particular problem f. Analyze similarities and differences among representations, solutions, or methods g. Prove that a solution is valid or that a method works for all similar cases h. Work on math problems that have multiple answers or solution methods i. Discuss math ideas, problems, solutions, or methods in small groups or pairs j. Work on mathematics textbook, worksheet, or board work exercises for practice or review k. Write extended explanations or math ideas, solutions, or methods l. Work on a mathematics investigation, problem, or project for several days 	<p><u>Response scale:</u></p> <ul style="list-style-type: none"> 1 = Never 2 = Once or twice a semester 3 = Once or twice a month 4 = Once or twice a week 5 = Almost every day



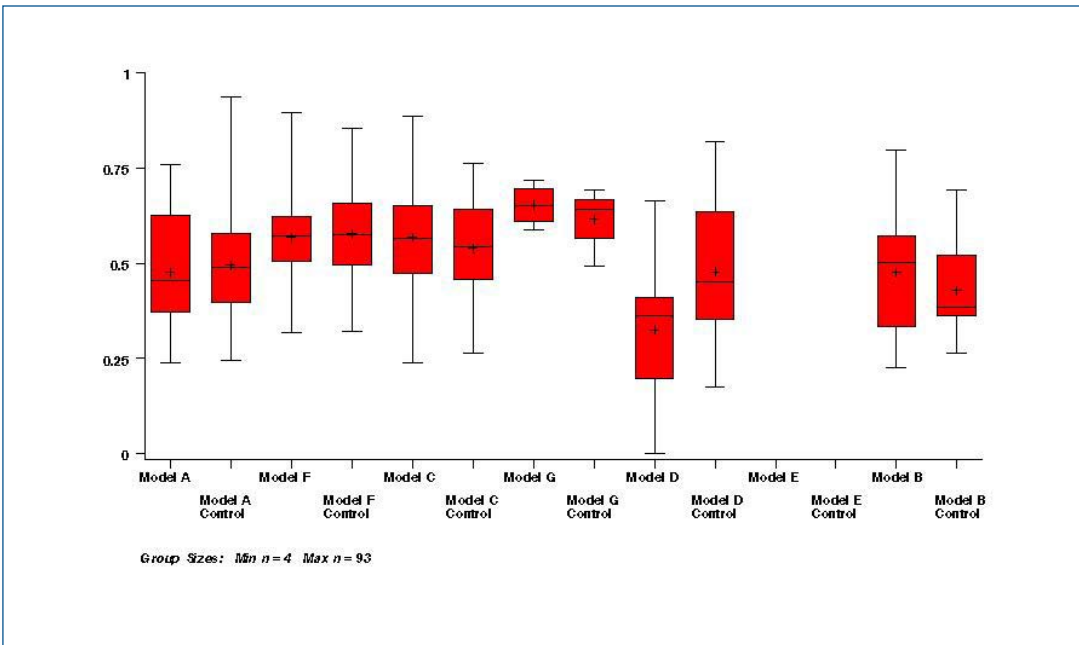
Appendix D: Box Plots Showing Variation and Implementation by CSR Key and by District, Year 1

Implementation by CSR Key

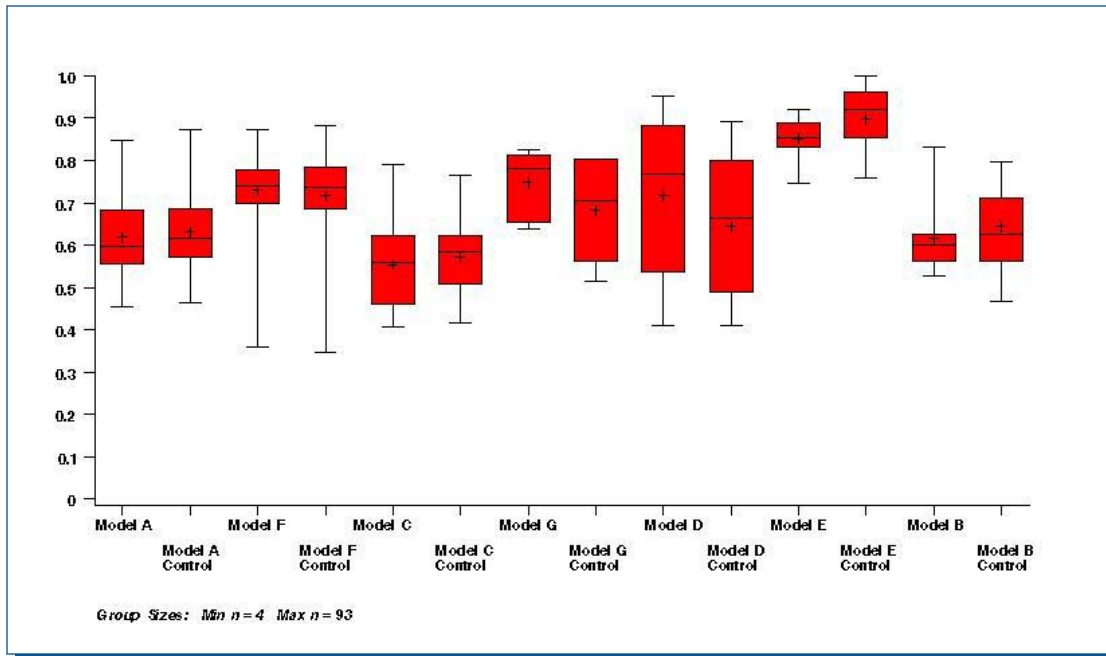
Implementation by CSR Key, Year 1: Shared Decision Making



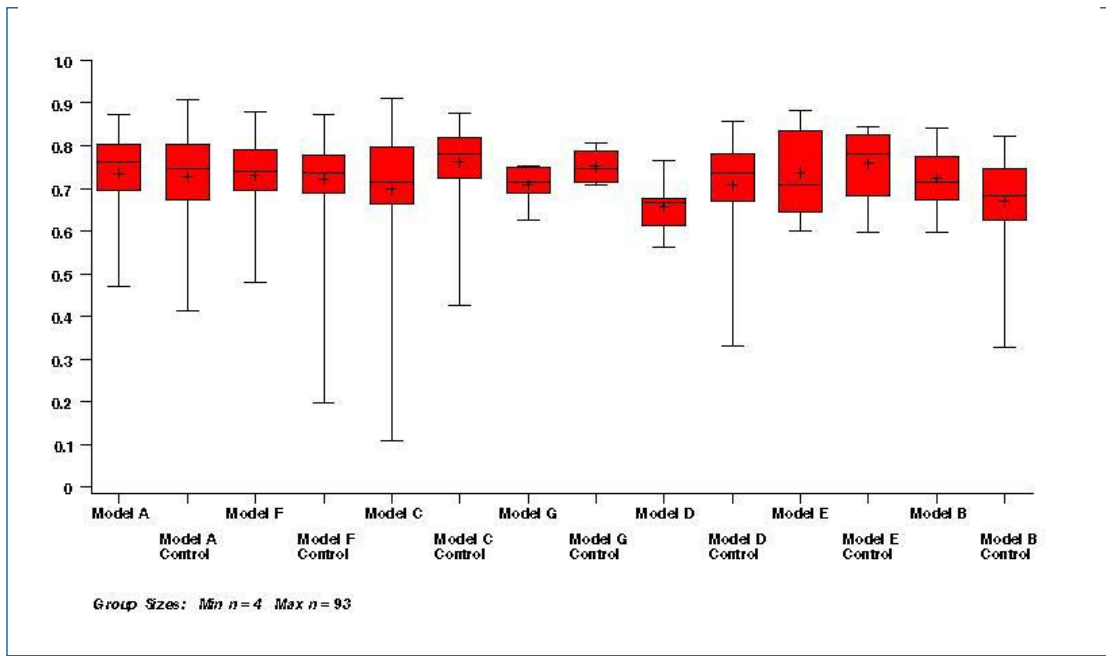
Implementation by CSR Key, Year 1: Use of Technology in Instruction



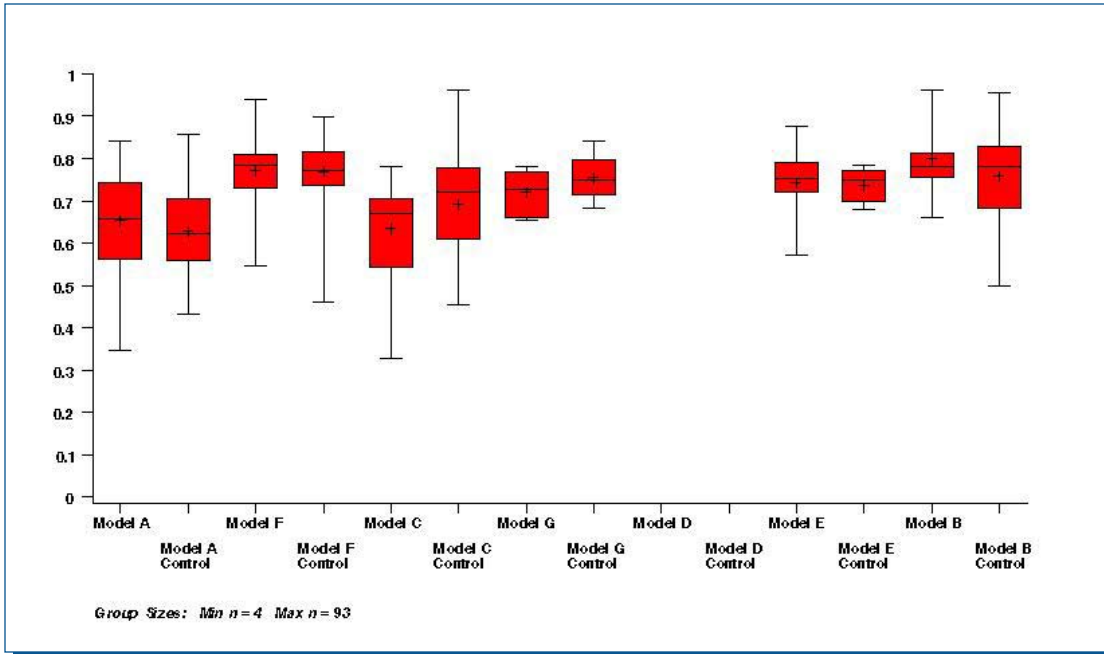
Implementation by CSR Key, Year 1: Parent/Community involvement



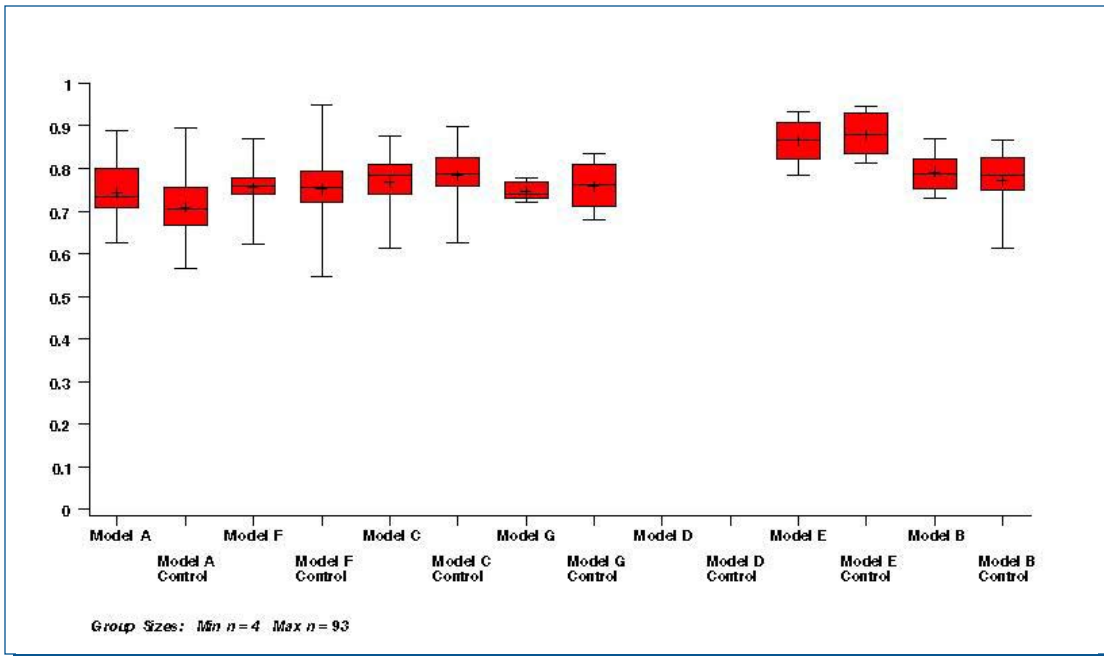
Implementation by CSR Key, Year 1: Emphasis of Professional Development



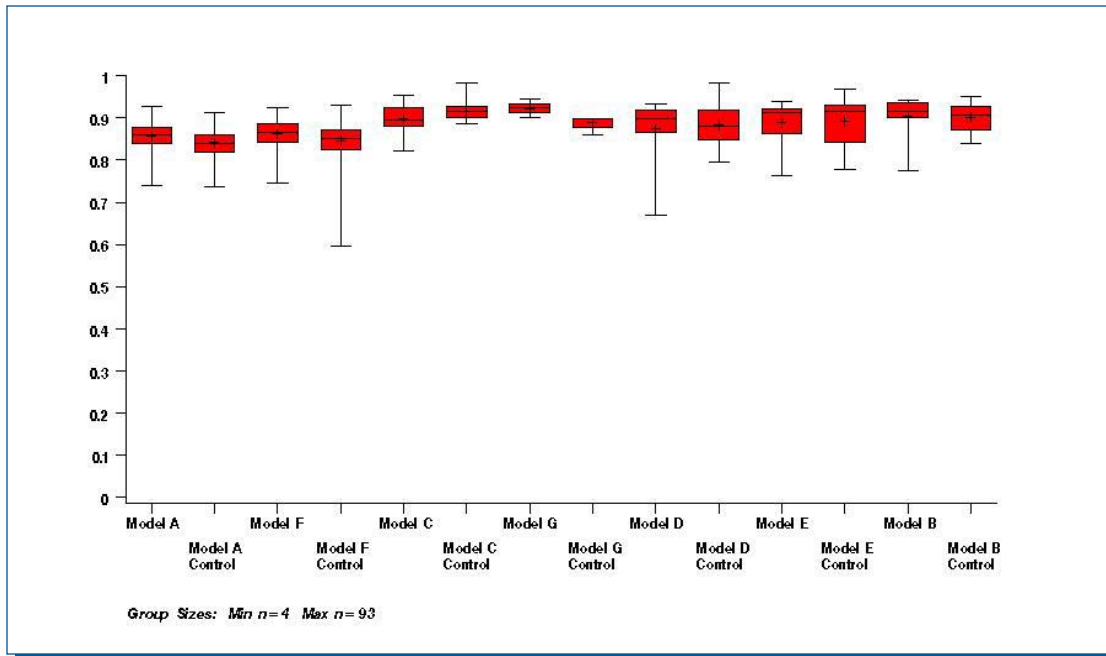
Implementation by CSR Key, Year 1: Engagement in Informal Professional Development



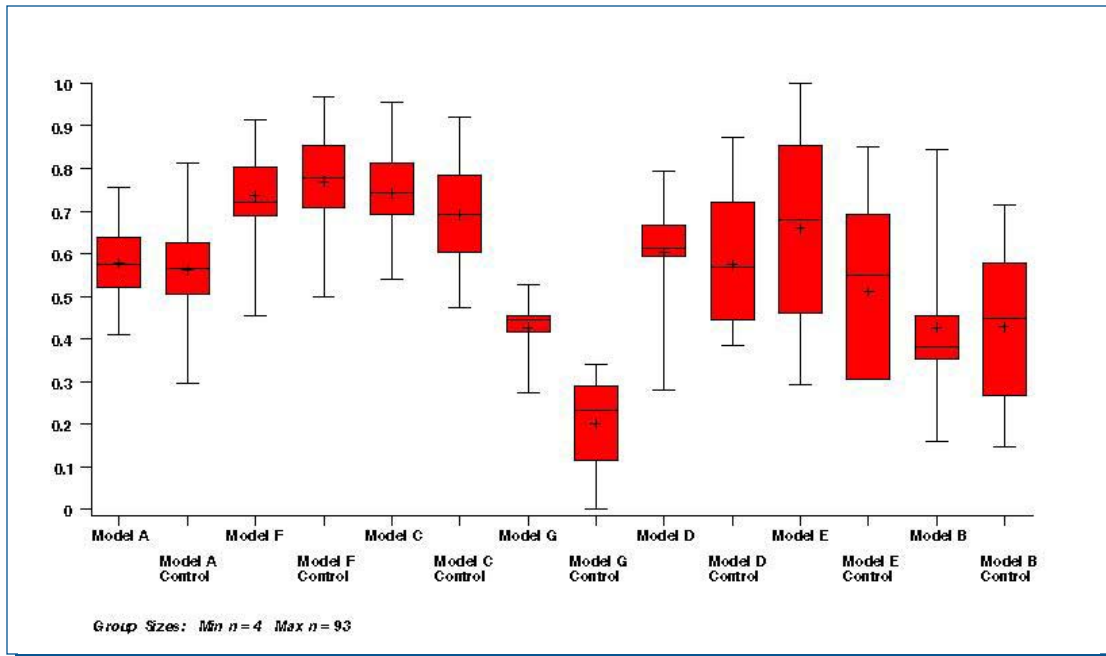
Implementation by CSR Key, Year 1: Influence of Assessments



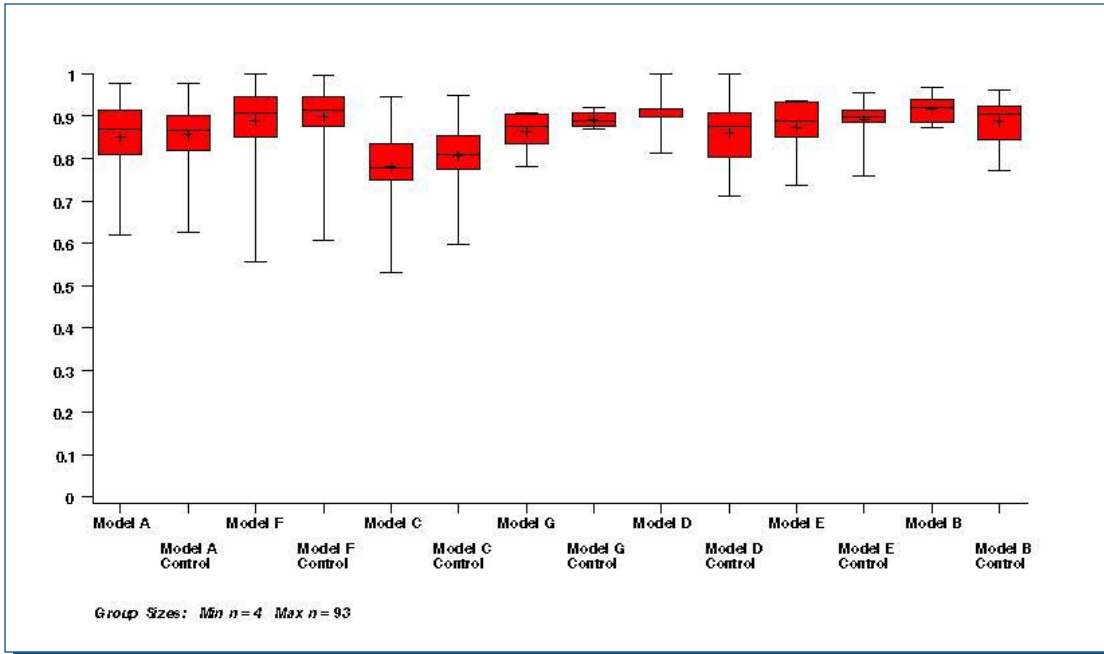
Implementation by CSR Key, Year 1: Use of Assessments



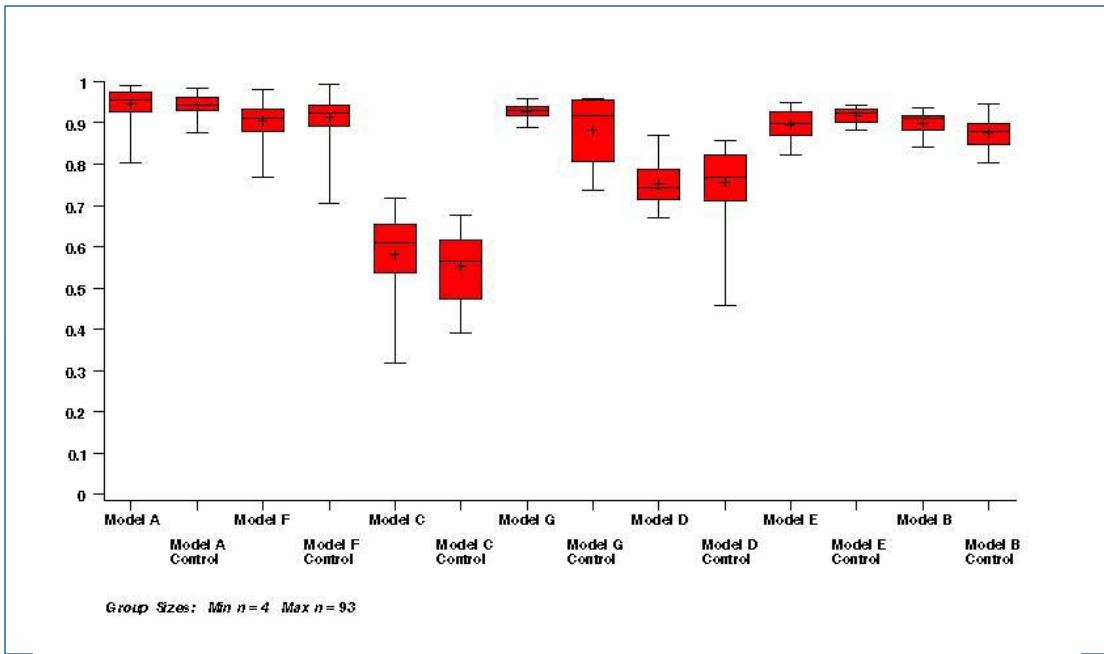
Implementation by CSR Key, Year 1: Inclusion



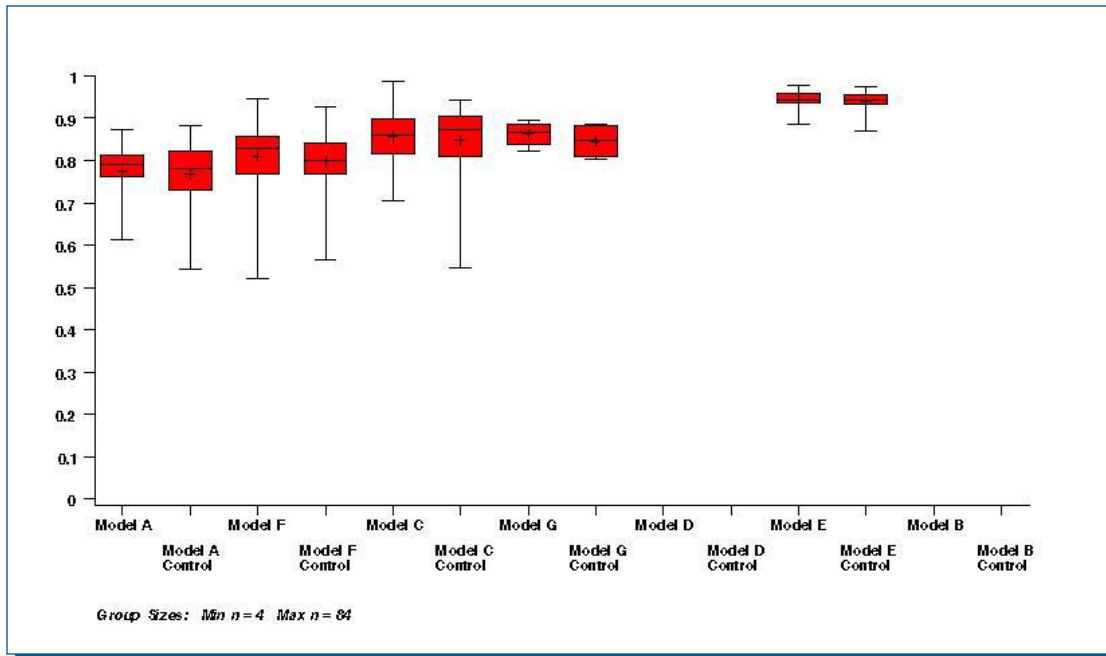
Implementation by CSR Key, Year 1: Student Grouping



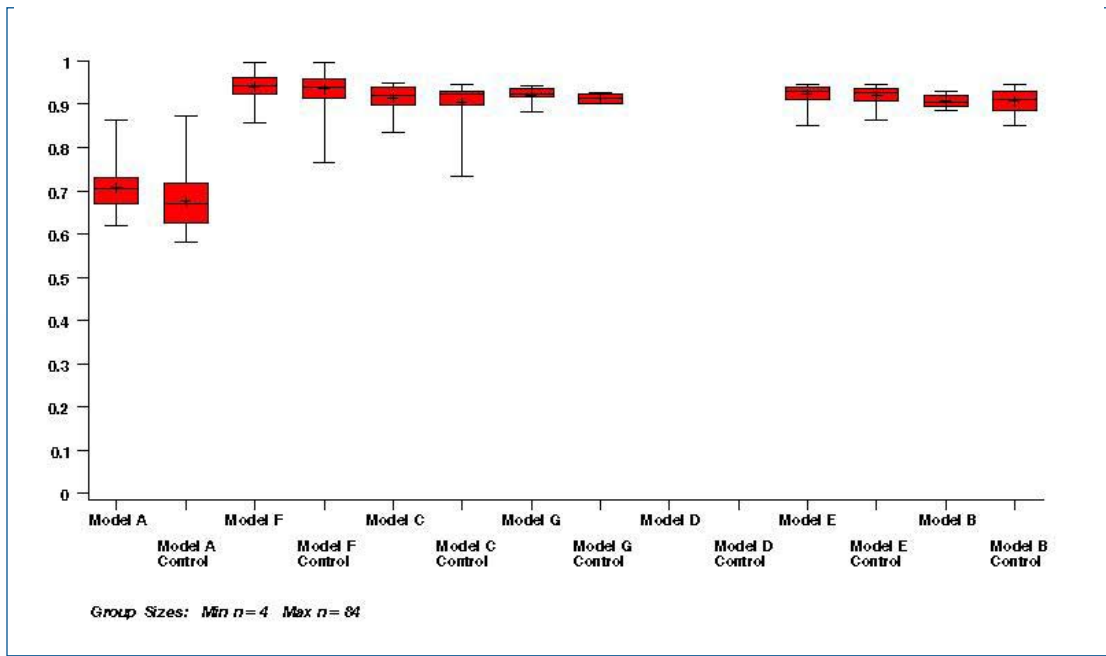
Implementation by CSR Key, Year 1: Time Scheduled for Instruction



Implementation by CSR Key, Year 1: Curriculum

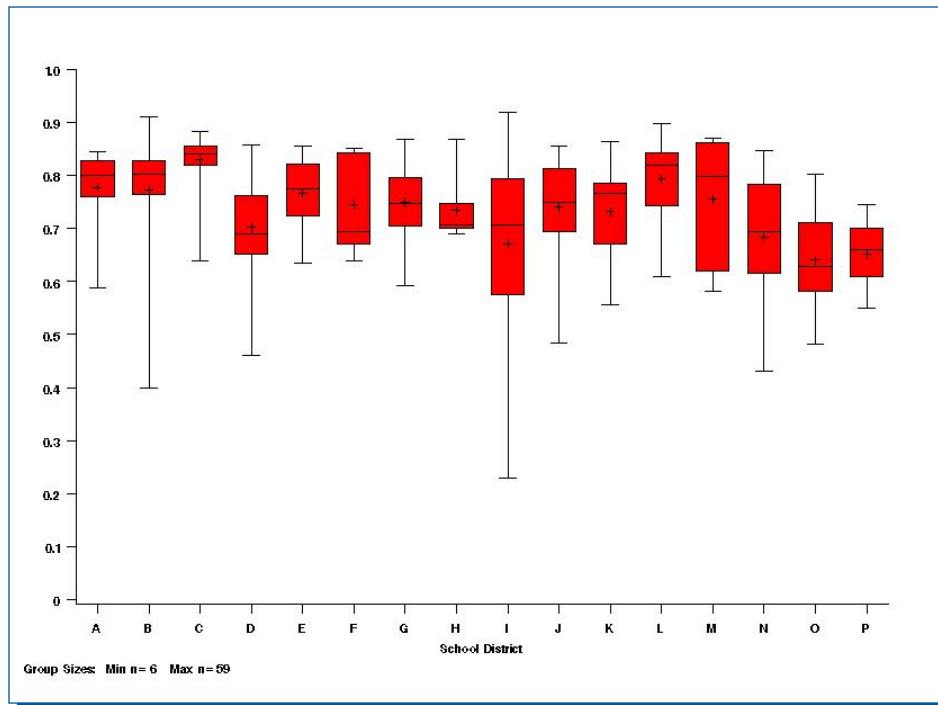


Implementation by CSR Key, Year 1: Pedagogy

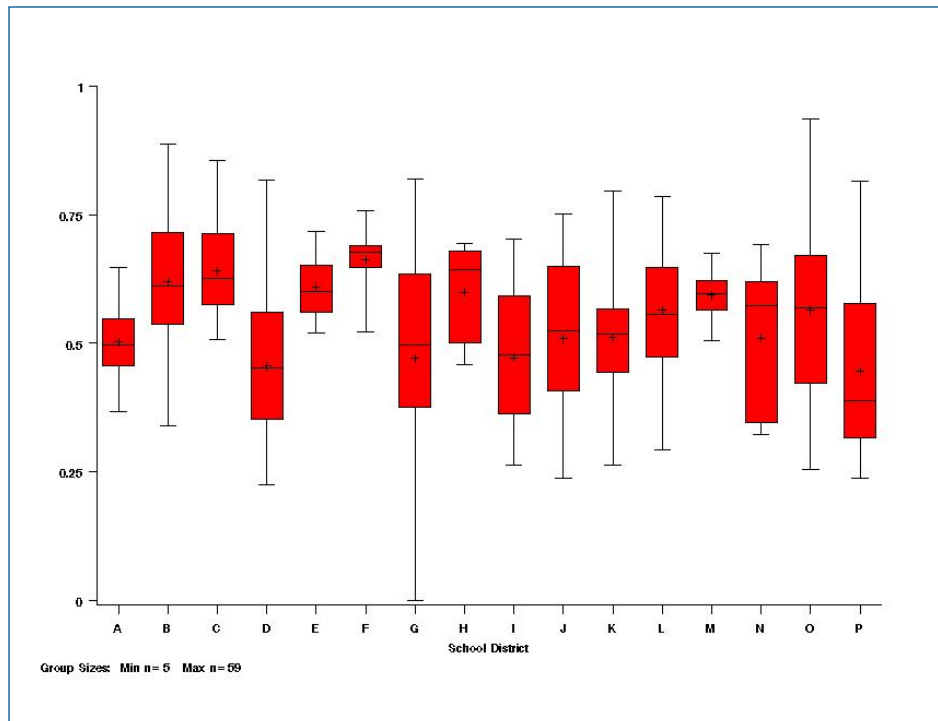


Implementation by District

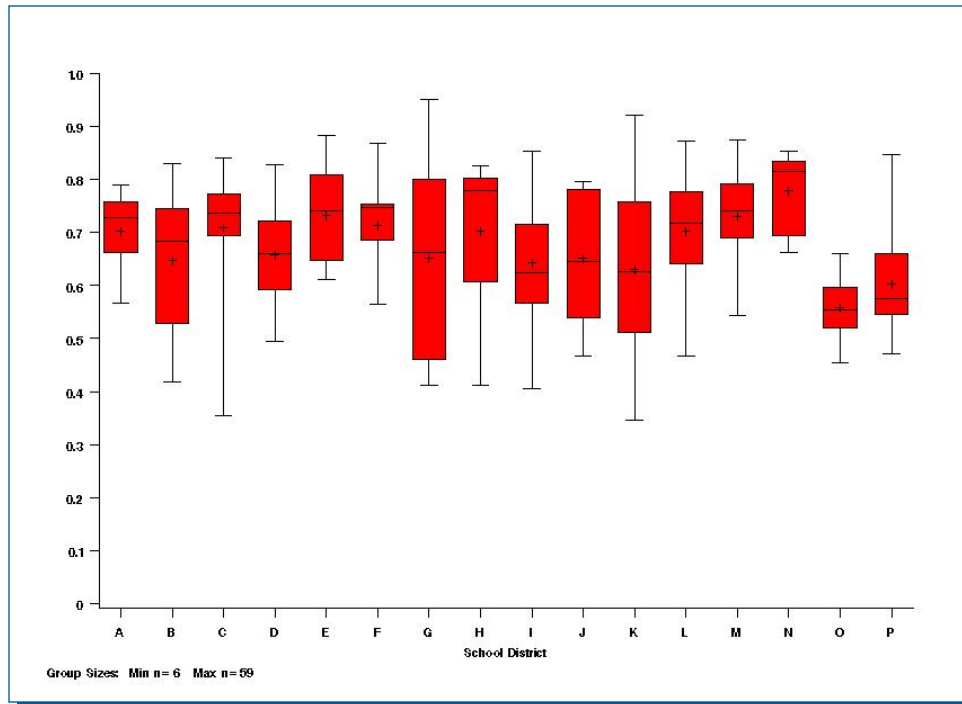
Implementation by District, Year 1: Shared Decision Making



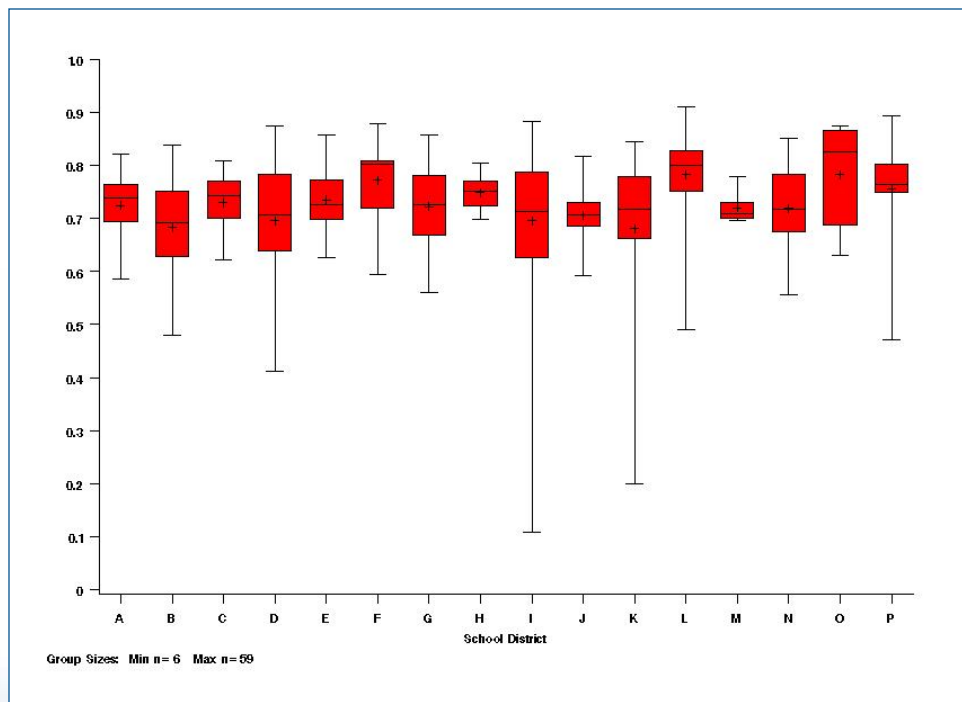
Implementation by District, Year 1: Use of Technology in Instruction



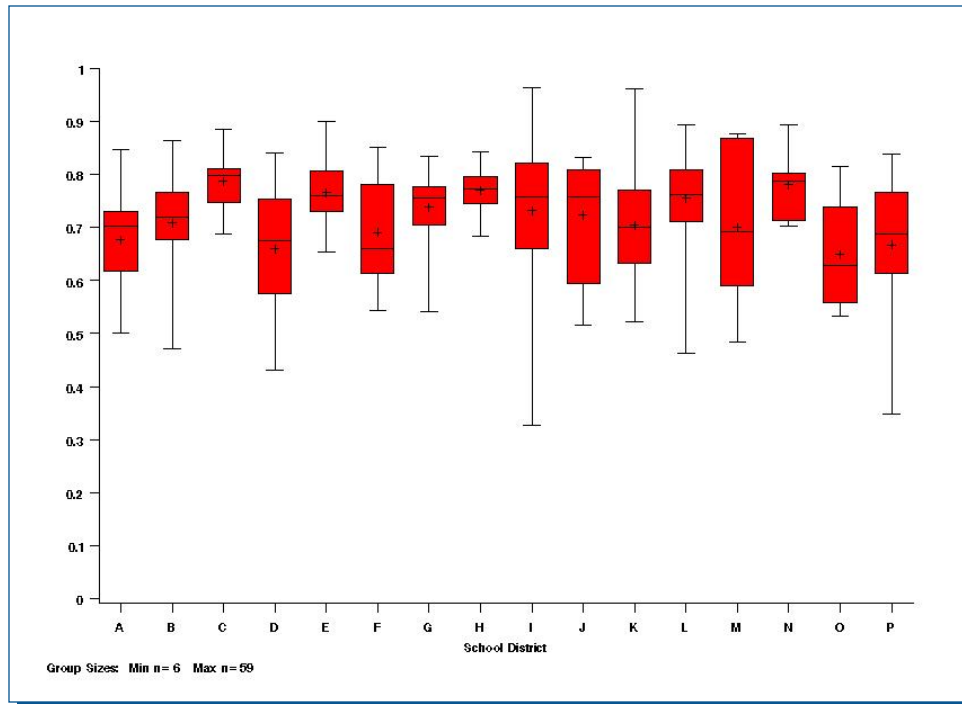
Implementation by District, Year 1: Parent/Community Involvement



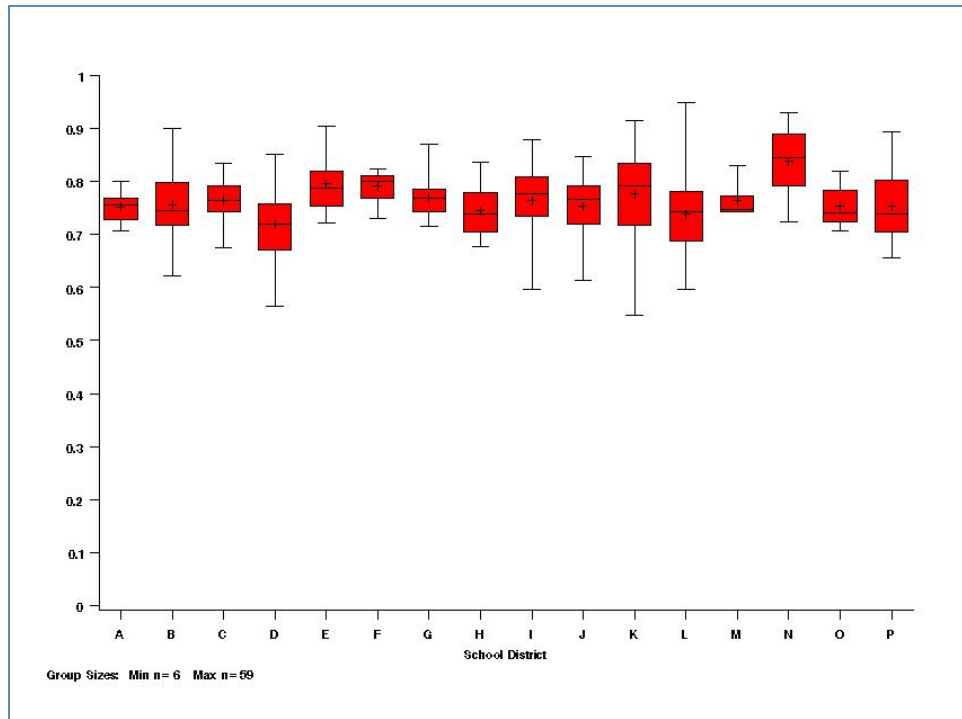
Implementation by District, Year 1: Emphasis of Professional Development



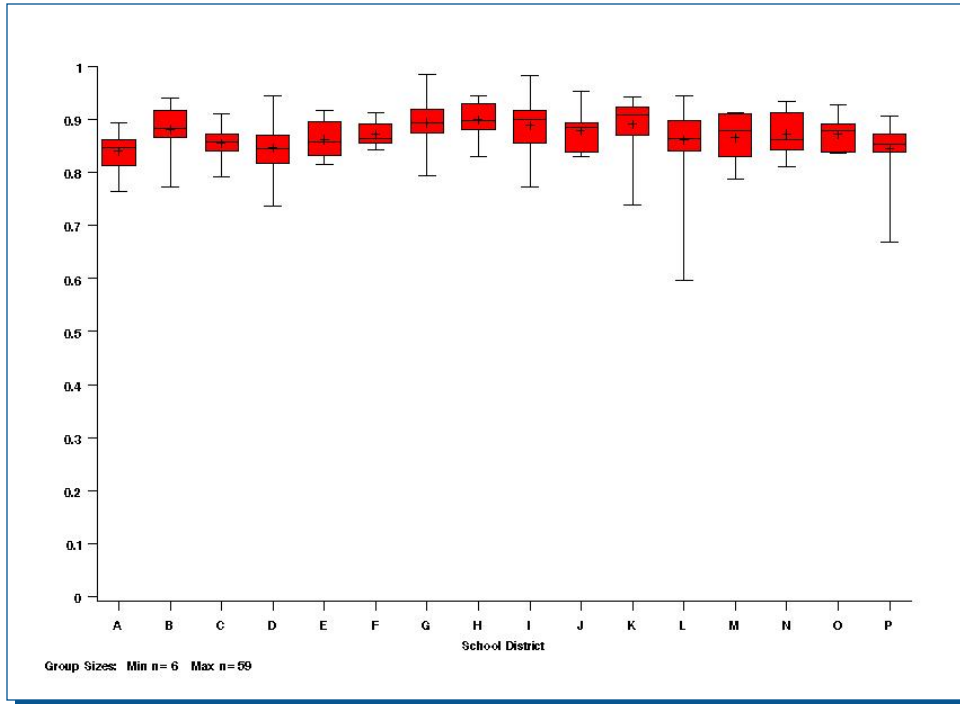
Implementation by District, Year 1: Engagement in Informal Professional Development



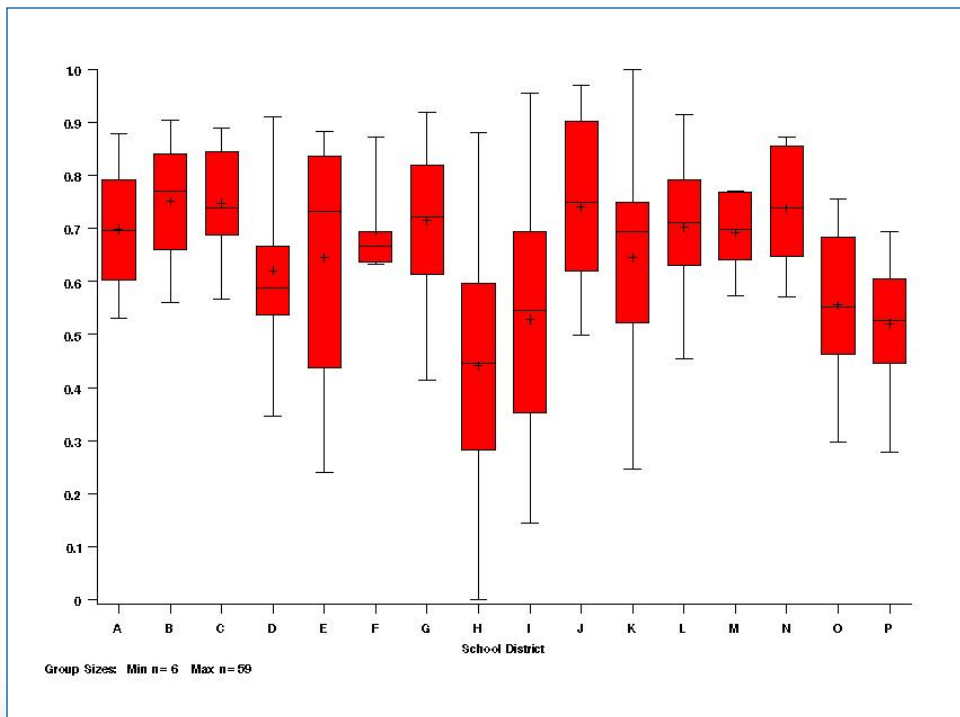
Implementation by District, Year 1: Influence of Assessments



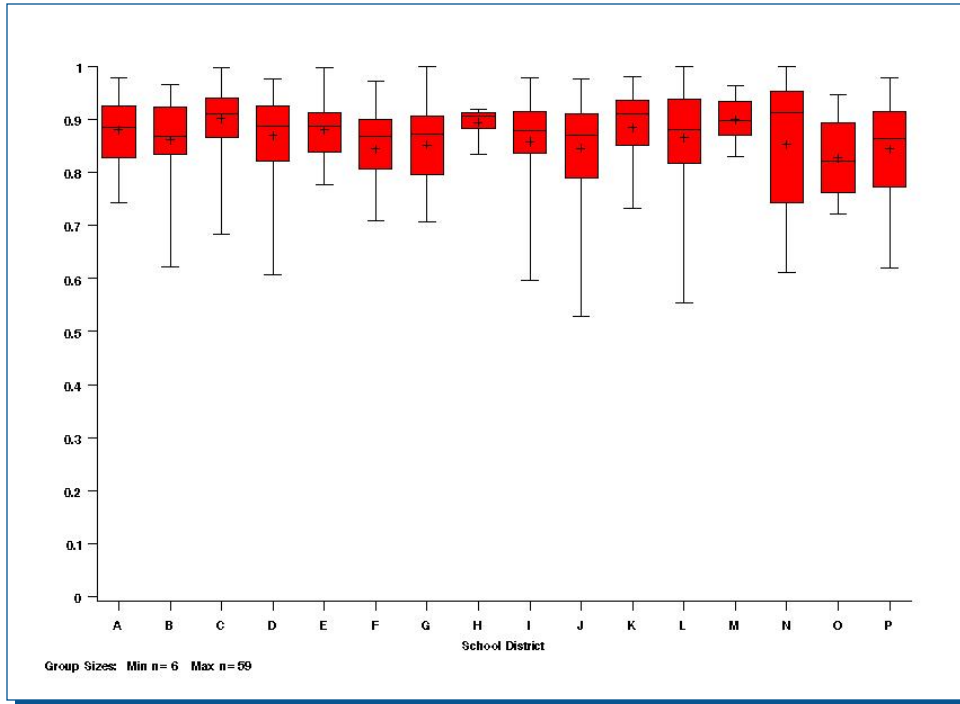
Implementation by District, Year 1: Use of Assessments



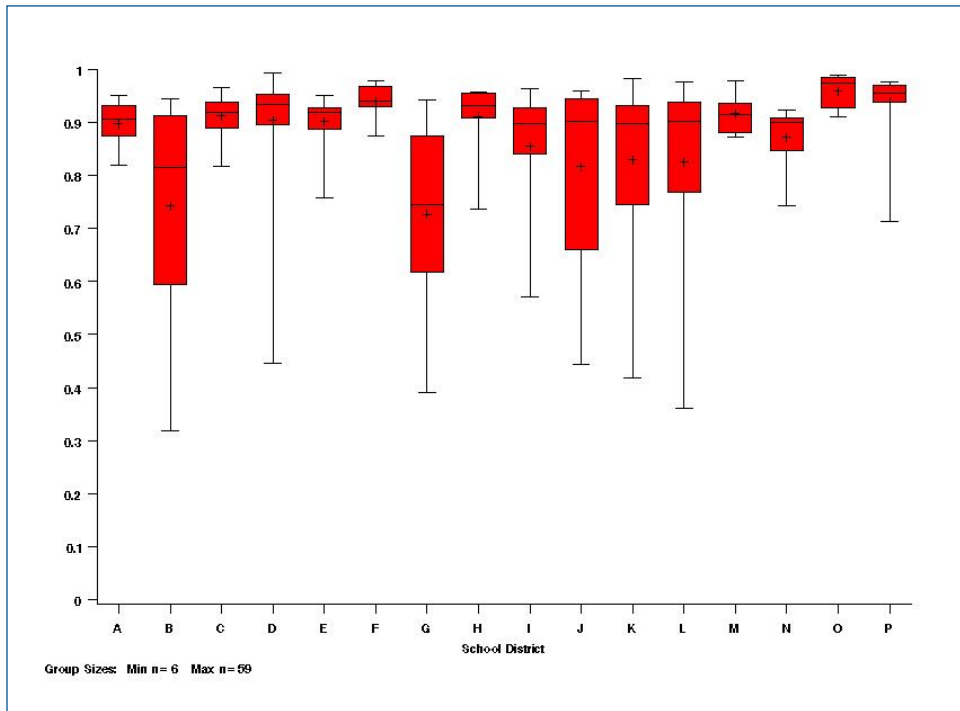
Implementation by District, Year 1: Inclusion



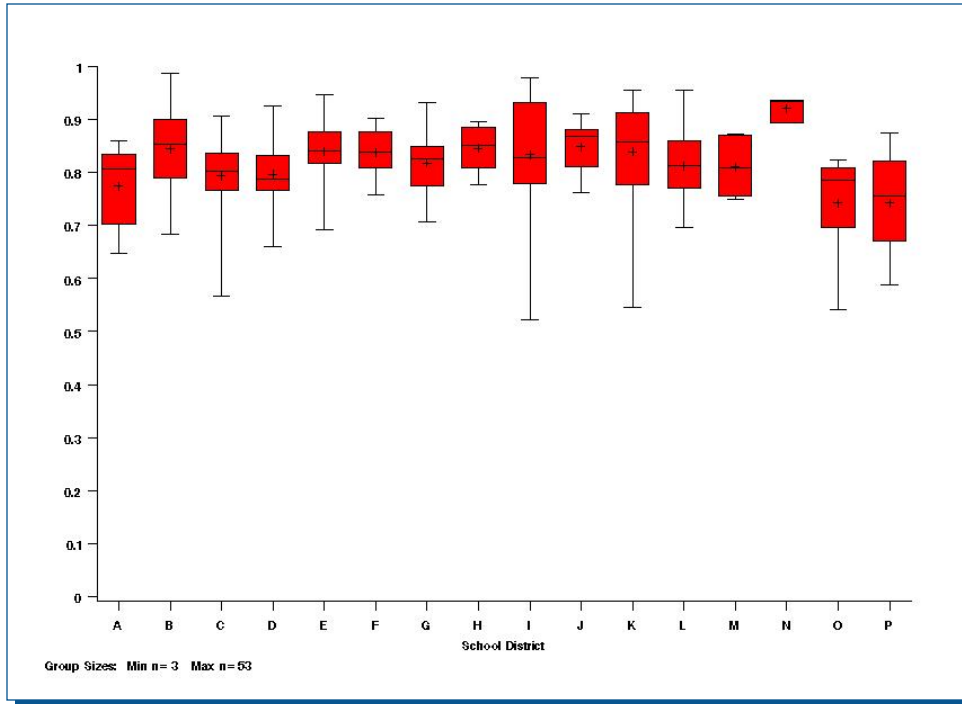
Implementation by District, Year 1: Student Grouping



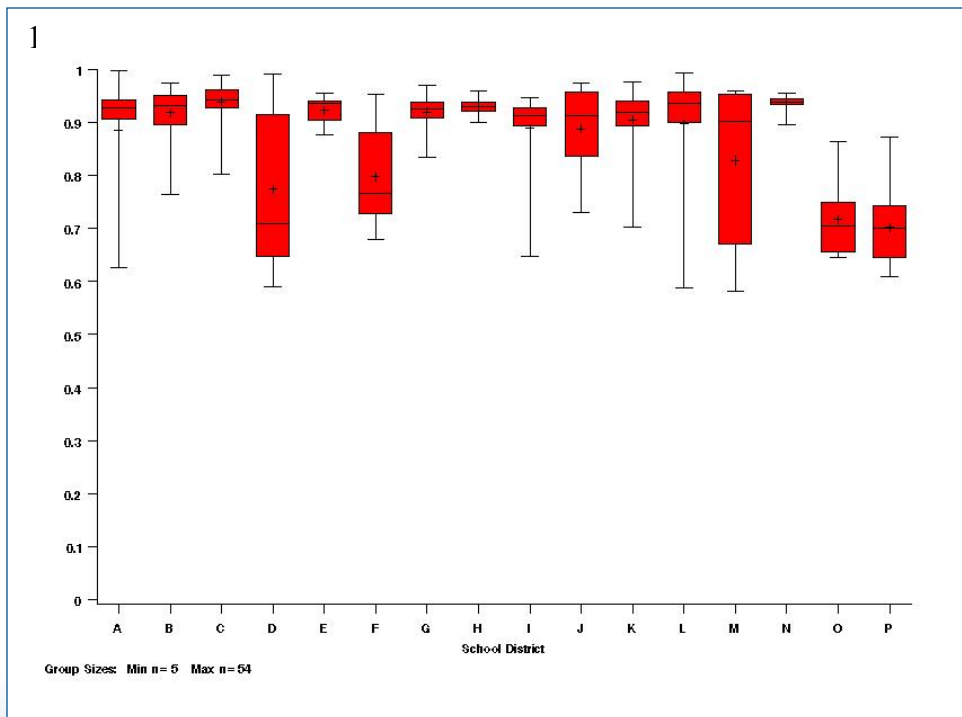
Implementation by District, Year 1: Time Scheduled for Instruction



Implementation by District, Year 1: Curriculum



Implementation by District, Year 1: Pedagogy



Appendix E: Significant Differences Among CSR Model Keys

Table E1. Significant Differences Among CSR Model Keys, Year 1

Shared Decision Making						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.65693336		0.1125	<.0001***	<.0001***	<.0001***
CSR Model B	0.62794660	0.1125		<.0001***	<.0001***	<.0001***
CSR Model C	0.76863889	<.0001***	<.0001***		<.0001***	0.9999
CSR Model F	0.82325348	<.0001***	<.0001***	<.0001***		0.0008***
CSR Model E	0.77070781	<.0001***	<.0001***	0.9999	0.0008***	
School Organization						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.66395249		0.9480	0.0020***	0.0666*	0.3711
CSR Model B	0.63556228	0.9480		0.0035*	0.0698*	0.2146
CSR Model C	0.77920745	0.0020***	0.0035***		0.2996	0.9513
CSR Model F	0.72786944	0.0666*	0.0698*	0.2996		0.9935
CSR Model E	0.74566732	0.3711	0.2146	0.9513	0.9935	
Use of Technology						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	
CSR Model A	0.49211236		0.7205	0.0186**	0.0002***	
CSR Model B	0.46107634	0.7205		0.0065***	0.0004***	
CSR Model C	0.56214983	0.0186**	0.0065***		0.9843	
CSR Model F	0.56952324	0.0002***	0.0004***	0.9843		
Parent/Community Involvement						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.63764455		0.9996	<.0001***	<.0001***	<.0001***
CSR Model B	0.63379481	0.9996		0.0030***	<.0001***	<.0001***
CSR Model C	0.56448112	<.0001***	0.0030***		<.0001***	<.0001***
CSR Model F	0.74368944	<.0001***	<.0001***	<.0001***		<.0001***
CSR Model E	0.87535536	<.0001***	<.0001***	<.0001***	<.0001***	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E1. Significant Differences Among CSR Model Keys, Year 1 (continued)

Emphasis of Professional Development						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.72632007		0.7730	0.8633	0.9978	1.0000
CSR Model B	0.70301602	0.7730		0.3347	0.5702	0.9406
CSR Model C	0.74212922	0.8633	0.3347		0.9151	0.9410
CSR Model F	0.73033042	0.9978	0.5702	0.9151		0.9982
CSR Model E	0.72365787	1.0000	0.9406	0.9410	0.9982	
Engagement in Informal Professional Development						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.63649975		<.0001***	0.2606	<.0001***	0.0010***
CSR Model B	0.78354731	<.0001***		<.0001***	0.9168	0.2575
CSR Model C	0.66895808	0.2606	<.0001***		<.0001***	0.0988*
CSR Model F	0.76807995	<.0001***	0.9168	<.0001***		0.3894
CSR Model E	0.72891419	0.0010***	0.2575	0.0988*	0.3894	
Influence of Assessments						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.72279978		<.0001***	<.0001***	0.0007***	<.0001***
CSR Model B	0.78415348	<.0001***		0.9291	0.1137	0.0002***
CSR Model C	0.77325851	<.0001***	0.9291		0.2863	<.0001***
CSR Model F	0.75550674	0.0007***	0.1137	0.2863		<.0001***
CSR Model E	0.86058721	<.0001***	0.0002***	<.0001***	<.0001***	
Use of Assessments						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.84810888		<.0001***	<.0001***	0.6012	0.0012***
CSR Model B	0.90126414	<.0001***		0.8854	<.0001***	0.5338
CSR Model C	0.90903883	<.0001***	0.8854		<.0001***	0.0877*
CSR Model F	0.85538780	0.6012	<.0001***	<.0001***		0.0094***
CSR Model E	0.88455209	0.0012***	0.5338	0.0877*	0.0094***	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E1. Significant Differences Among CSR Model Keys, Year 1 (continued)

Inclusion						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.57625736		<.0001***	<.0001***	<.0001***	0.9979
CSR Model B	0.44434383	<.0001***		<.0001***	<.0001***	0.0009***
CSR Model C	0.70869804	<.0001***	<.0001***		0.3404	0.0014***
CSR Model F	0.74270498	<.0001***	<.0001***	0.3404		<.0001***
CSR Model E	0.58606283	0.9979	0.0009***	0.0014***	<.0001***	
Student Grouping						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.86080258		0.0894*	<.0001***	0.0226	0.9605
CSR Model B	0.90498988	0.0894*		<.0001***	0.9618	0.7057
CSR Model C	0.78832061	<.0001***	<.0001***		<.0001***	0.0005***
CSR Model F	0.89416086	0.0226**	0.9618	<.0001***		0.8540
CSR Model E	0.87480375	0.9605	0.7057	0.0005***	0.8540	
Time for Teaching						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.94310542		<.0001***	<.0001***	<.0001***	0.0752*
CSR Model B	0.88740692	<.0001***		<.0001***	0.3550	0.6782
CSR Model C	0.57056106	<.0001***	<.0001***		<.0001***	<.0001***
CSR Model F	0.90684012	<.0001***	0.3550	<.0001***		1.0000
CSR Model E	0.90800774	0.0752*	0.6782	<.0001***	1.0000	
Curriculum						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>	
CSR Model A	0.77158508		<.0001***	0.0043***	<.0001***	
CSR Model C	0.85307025	<.0001***		<.0001***	<.0001***	
CSR Model F	0.80497575	0.0043***	<.0001***		<.0001***	
CSR Model E	0.93868414	<.0001***	<.0001***	<.0001***		

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E1. Significant Differences Among CSR Model Keys, Year 1 (continued)

Pedagogy						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.69225474		<.0001***	<.0001***	<.0001***	<.0001***
CSR Model B	0.90579855	<.0001***		0.9722	0.0010***	0.9718
CSR Model C	0.91155267	<.0001***	0.9722		0.0003***	0.9999
CSR Model F	0.93816274	<.0001***	0.0010***	0.0003***		0.0880*
CSR Model E	0.91326687	<.0001***	0.9718	0.9999	0.0880*	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

Table E2. Significant Differences Among CSR Model Keys, Year 3

Shared Decision-Making						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.74072381		<.0001***	0.8106	0.2033	0.9559
CSR Model B	0.63329410	<.0001***		0.0005***	<.0001***	0.0002***
CSR Model C	0.72211243	0.8106	0.0005***		0.0204**	0.6470
CSR Model F	0.76839677	0.2033	<.0001***	0.0204**		0.9941
CSR Model E	0.75853236	0.9559	0.0002***	0.6470	0.9941	0.9559
School Organization						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.56973489			<.0001***	1.0000	0.0011***
CSR Model B	Non-est	.				
CSR Model C	0.72279547	<.0001***			<.0001***	0.8770
CSR Model F	0.56861148	1.0000		<.0001***		0.0005***
CSR Model E	0.76275728	0.0011***		0.8770	0.0005***	
Use of Technology						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	
CSR Model A	0.55886648		0.4176	0.9608	0.7950	
CSR Model B	0.49853569	0.4176		0.2737	0.0988*	
CSR Model C	0.57555569	0.9608	0.2737		0.9965	
CSR Model F	0.58205224	0.7950	0.0988*	0.9965		

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E2. Significant Differences Among CSR Model Keys, Year 3 (continued)

Parent/Community Involvement						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.63562861		<.0001***	0.0010***	<.0001***	<.0001***
CSR Model B	0.73369974	<.0001***		<.0001***	0.0234**	0.0038***
CSR Model C	0.56763720	0.0010***	<.0001***		<.0001***	<.0001***
CSR Model F	0.78984705	<.0001***	0.0234**	<.0001***		0.3045
CSR Model E	0.83696325	<.0001***	0.0038***	<.0001***	0.3045	
Emphasis of Professional Development						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.76540440		0.9327	0.9124	0.5284	0.8461
CSR Model B	0.74101789	0.9327		1.0000	0.9994	0.9946
CSR Model C	0.74264784	0.9124	1.0000		0.9972	0.9908
CSR Model F	0.73466783	0.5284	0.9994	0.9972		0.9981
CSR Model E	0.72230382	0.8461	0.9946	0.9908	0.9981	
Engagement in Informal Professional Development						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.65859847		<.0001***	0.0026***	<.0001***	0.0333**
CSR Model B	0.81000987	<.0001***		0.1913	0.9996	0.8470
CSR Model C	0.74570308	0.0026***	0.1913		0.0433**	0.9810
CSR Model F	0.80492486	<.0001***	0.9996	0.0433**		0.8408
CSR Model E	0.76766489	0.0333**	0.8470	0.9810	0.8408	
Influence of Assessments						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.70450138		<.0001***	<.0001***	<.0001***	0.3801
CSR Model B	0.79288943	<.0001***		1.0000	0.7621	0.6286
CSR Model C	0.79194216	<.0001***	1.0000		0.6615	0.5967
CSR Model F	0.77195603	<.0001***	0.7621	0.6615		0.9325
CSR Model E	0.75186745	0.3801	0.6286	0.5967	0.9325	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E2. Significant Differences Among CSR Model Keys, Year 3 (continued)

Use of Assessments						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.83722264		<.0001***	<.0001***	0.4842	0.0133**
CSR Model B	0.90751836	<.0001***		0.9999	0.0004***	0.9978
CSR Model C	0.90946264	<.0001***	0.9999		<.0001***	0.9929
CSR Model F	0.85226250	0.4842	0.0004***	<.0001***		0.0842*
CSR Model E	0.90057586	0.0133**	0.9978	0.9929	0.0842*	
Inclusion						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.75030725		0.0099***	0.1395	<.0001***	0.0747*
CSR Model B	0.66950178	0.0099***		<.0001***	<.0001***	<.0001***
CSR Model C	0.79693278	0.1395	<.0001***		0.6820	0.8636
CSR Model F	0.82032389	<.0001***	<.0001***	0.6820		0.9990
CSR Model E	0.82758650	0.0747*	<.0001***	0.8636	0.9990	
Student Grouping						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.87764346		0.0511*	0.0185**	<.0001***	0.8969
CSR Model B	0.93215334	0.0511*		<.0001***	0.9995	0.8611
CSR Model C	0.82448061	0.0185**	<.0001***		<.0001***	0.0541 *
CSR Model F	0.93623032	<.0001***	0.9995	<.0001***		0.6934
CSR Model E	0.90223224	0.8969	0.8611	0.0541*	0.6934	
Time for Teaching						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.94250137		0.0485**	<.0001***	0.2380	0.4268
CSR Model B	0.90120525	0.0485**		<.0001***	0.5378	0.9992
CSR Model C	0.48515588	<.0001***	<.0001***		<.0001***	<.0001***
CSR Model F	0.92238444	0.2380	0.5378	<.0001***		0.9363
CSR Model E	0.90682861	0.4268	0.9992	<.0001***	0.9363	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$



Table E2. Significant Differences Among CSR Model Keys, Year 3 (continued)

Curriculum						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>	
CSR Model A	0.80054943		0.0023***	0.9606	<.0001***	
CSR Model C	0.86007999	0.0023***		0.0001***	0.0638*	
CSR Model F	0.79424922	0.9606	0.0001***		<.0001***	
CSR Model E	0.94228147	<.0001***	0.0638*	<.0001***		
Pedagogy						
	<i>Imp Mean</i>	<i>CSR Model A</i>	<i>CSR Model B</i>	<i>CSR Model C</i>	<i>CSR Model F</i>	<i>CSR Model E</i>
CSR Model A	0.68957950		<.0001***	<.0001***	<.0001***	<.0001***
CSR Model B	0.91700975	<.0001***		0.8603	0.0418**	0.6782
CSR Model C	0.93234405	<.0001***	0.8603		0.2946	0.9659
CSR Model F	0.95456660	<.0001***	0.0418 **	0.2946		0.9913
CSR Model E	0.94588652	<.0001***	0.6782	0.9659	0.9913	

Note. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$

