

## **The Construct Validity of Teachers' Perceptions of Change in Schools Implementing Comprehensive School Reform Models**

*John A. Nunnery  
Old Dominion University*

*Steven M. Ross  
The University of Memphis*

*Linda Bol  
Old Dominion University*

### *Abstract*

*This study reports the results of a validation study of the Comprehensive School Restructuring Teacher Questionnaire (CSRTQ) and the School Observation Measure (SOM), which are intended for use in evaluating comprehensive school reform efforts. The CSRTQ, which putatively measures five factors related to school restructuring (internal focus, external support, school capacity, pedagogical change, and restructuring outcomes), was administered to 2,511 teachers in 83 schools that received Comprehensive School Reform grants. The SOM, a high-inference observation instrument that measures school-wide implementation of various instructional practices, was administered eight to ten times in each of the 83 schools. Confirmatory factor analysis supported the a priori factor structure of the CSRTQ. Hierarchical linear modeling analysis indicated that teacher perceptions of pedagogical change were significantly predictive of aggregate SOM observations of specific teaching practices, and that the SOM items reliably distinguish between schools on 26 teaching practices. Mean pedagogical change scores were negatively related to traditional, teacher-centered practices, and positively related to the student-centered practices advanced by most CSR restructuring designs.*

## Introduction

Spurred by changes in federal policy and the New American Schools initiative (Borman, Hewes, Overman, & Brown, 2003), an unprecedented number of schools in the United States are simultaneously undertaking the implementation of whole school reforms with the assistance of external “design based assistance organizations” (Bodilly, 1996). These design-based assistance organizations (DBAO) provide training, materials, and technical assistance to support whole school reform (Bodilly, 2001). Choosing a DBAO with which to pursue restructuring is hardly as simple as choosing a model that “works.” The success of any design-based restructuring effort is not only contingent on an appropriate match between characteristics of the design and the needs of the school, but also on the complex interplay of school capacity, quality and intensity of external support, and local history. Despite the provision in the Comprehensive School Reform (CSR; U.S. Department of Education, 2007) legislation that schools adopt “research-based” school restructuring designs, further research is needed to determine whether the adoption of designs results in substantial changes in schools’ curricula, use of research-based instructional strategies, parental involvement, or other organizational features that are the focus of Title 1 program priorities (Puma & Drury, 2000).

Recent research on comprehensive school reform has been conducted from a variety of perspectives, from case studies of individual schools, to comprehensive evaluations of scale-up efforts in large districts, to national evaluations of design-based restructuring efforts (e.g., Bodilly, 2001; Datnow, 2000a; Ross et al., 1997). Likewise, a multiplicity of research methodologies have been employed to examine implementation and outcomes, including case study analysis, observation of instruction, interviews, questionnaire surveys of teachers, and analysis of student achievement data (e.g., Ross et al., 2001; Stringfield, Datnow, Ross, & Snively, 1998). Although the diversity of research in this area may be regarded as a strength, it is also problematic that program implementation data tend to be specific to individual designs in the extant literature on comprehensive school restructuring.

Thus, it is difficult to examine the relative effectiveness of design-based assistance organizations in promoting organizational change, or to examine important issues such as the interaction of design features and

school organizational variables on implementation of school restructuring designs. A recent meta-analysis of CSR effects found that, although students attending CSR schools significantly outperform students attending matched comparison schools, there is considerable unexplained variability in the achievement effects of CSR schools (Borman, Carter, Aladjem, & LeFloch, 2004).

This study examines the reliability and validity of two instruments that have been used to assess restructuring implementation and factors that affect implementation (e.g., Lewis, Ross, & Alberg, 1999; Ross, Smith, & Alberg, 1998; Ross, Smith, Alberg, & Lowther, 2004), which might then be used to explain some of the variability in achievement effects within and across schools implementing CSR designs. One instrument, the *School Observation Measure* (SOM; Ross, Smith, & Alberg, 1998), is a high inference observation instrument designed to measure school-wide use of a number of instructional strategies, including traditional practices (e.g., direct instruction and independent seatwork) and alternative, predominately student-centered methods associated with many educational reforms (e.g., cooperative learning, project-based learning, inquiry, discussion, using technology as a learning tool). The second instrument, the *Comprehensive School Reform Teacher Questionnaire* (CSRTQ; Ross & Alberg, 1999), solicits teachers' perceptions on items related to five constructs often used to describe and interpret findings in the school restructuring literature.

### **Constructs Underlying Teacher Perceptions of Comprehensive School Restructuring: A Five-Factor Model**

Despite the diversity of contexts, perspectives, and methods from which knowledge has been produced, analysis of comprehensive school restructuring processes and outcomes tends to revolve around a relatively small number of broad constructs. Based on a comprehensive, longitudinal research program on school restructuring, Ross, Stringfield, Nunnery and colleagues have identified five constructs that repeatedly emerge in the literature and in their own empirical studies (e.g., Nunnery, 1998; Ross et al., 2000; Ross et al., 1997; Stringfield & Datnow, 1998; Stringfield, Waxman, & Padron, 2000). These five constructs are school capacity, external support, internal focus, pedagogical change, and restructuring outcomes.

*School capacity.* School capacity indicates whether the school has sufficient resources to implement a restructuring effort. Several studies highlight the influence that school capacity variables have in restructuring efforts. The timely availability of appropriate materials and technology is a crucial prerequisite to change in teaching practices (Datnow & Stringfield, 2000; Ross et al., 2000). Similarly, sufficient staffing and planning time must be allocated to achieve focus and change (Bodilly, 2001; Bodilly, 1996; Bol et al., 1998; Stringfield et al., 1998). Schools and districts must also ensure that adequate fiscal resources are available to support acquisition of staff, materials, and technical assistance (Glennan, 1998; Ross et al., 1997).

*External support.* External support refers to the quality and quantity of assistance and direction a school receives from outside the immediate organization (e.g., from the school district, state, or DBAO) to implement restructuring. The importance of external support for school restructuring is well-documented (e.g., Crandall & Loucks, 1983; Desimone, 2002; Fullan, 1991; Hopkins, Ainscow, & West, 1994; McLaughlin, 1990). Within the context of research on implementation of comprehensive school restructuring models, a number of important external supports have been identified. Several studies point to the importance of the quality of professional development provided by the external design team (Bodilly, 1996; Bol et al., 1998; Nunnery et al., 1997; Ross et al., 2000). Technical assistance and coordination provided by district and state personnel also seem to facilitate successful implementation of restructuring models (Bodilly, 2001; Datnow, 2000a; Smith et al., 1998; Stringfield et al., 1998). Both the quantity and quality of professional development and technical assistance are predictive of teachers' understanding of the practices embedded in the restructuring design adopted by their school. Not surprisingly, the CSR program requires use of high quality external support and assistance from an entity that has experience and expertise in school-wide reform and improvement (U.S. Department of Education, 2007).

As supported by CSR research over the last several years, schools' and DBAOs' failure to integrate external support, particularly from the local districts, with their reform activities is arguably one of the key factors that limited success and sustainability in many cases (Borman et al., 2004; Ross & Gil, 2004; Rowan, Camburn, & Barnes, 2004). McLaughlin and Talbert (2003), for example, identified five district activities as predictive of successful partnerships between districts and CSR schools: (a) taking a systems approach

to reform by seeing themselves (the districts) as the unit of change, (b) creating a learning community in the district office, (c) focusing resource allocation on teaching and learning, (d) supporting schools' professional learning and instructional improvement needs, and (e) using data to support accountability by the district and schools. Clearly, external support is an essential ingredient in the effective implementation of CSR models.

*Internal focus.* Focus refers to the degree to which the comprehensive school reform effort becomes central to the activities of school personnel. Centrality of a reform to the mission of the school has been found to play an important role in the adoption of educational innovations and in producing change in teaching practices (Berman & McLaughlin, 1977; Bodilly, 1996; Havelock, 1971). Several factors contribute to focus, and have been shown to be predictive of pedagogical change and continuation of reforms. Teacher buy-in and support of the change consistently predicts effectiveness and longevity of school reform efforts (Cooper, Slavin, & Madden, 1998; Datnow, 2000b; Ross et al., 2000; Rowan et al., 2004; Smith et al., 1997). Coordination and alignment of external resources and mandates also seems to influence the degree to which the reform becomes the focus for the activities of school personnel (Ross et al., 1997; Rowan et al., 2004; Stringfield & Datnow, 1998). Internally, the integration of the restructuring model with ongoing school efforts, teaching practices, and other contextual variables may also influence the pace and depth of change, as well as teacher support of the change effort (Berends, Bodilly, & Kirby, 2002; Metz, 1998). Finally, formal attention given to benchmarking progress and evaluating change is both an indicator of focus and a predictor of implementation quality (Bodilly, 1996; Fuhrman, Clune, & Elmore, 1988; Ross, Alberg, & Nunnery, 1999; Ross & Gil, 2004; Ross et al., 1997; Rowan et al., 2004; Slavin, 1999).

*Pedagogical change.* Pedagogical change indicates the degree to which instructional practices change in a way commensurate with the goals of the restructuring model. School restructuring models vary greatly in terms of the form of pedagogy advocated, as well as the processes by which pedagogical change is putatively effected. Although the approaches may differ in substantive ways, schools implementing comprehensive school reform models typically seek to change several common elements of teaching and learning, such as reduced emphasis on instructional "drill" as represented by an emphasis on workbooks and worksheets, and increased emphasis on use

of technology, cooperative learning teams, and project-based work (Bodilly, 1996; Stringfield, Ross, & Smith, 1996). In a recent meta-analysis, Hamilton et al. (2003) found consistently positive effects linking instructional practices in the National Science Foundation Systemic Initiatives (SI) program and student achievement. A characteristic of this initiative is “an emphasis on instruction that engages students as active participants in their own learning and that enhances the development of complex cognitive skills and processes” (p. 3). These goals are realized through specific practices that include cooperative group learning, application of content to real world problems, inquiry-based activities, and open-ended assessments. Cohen, Raudenbush, and Ball (2003) argue that studies on the effectiveness of reform should focus on instruction or pedagogy as causal agents in student achievement gains.

*Restructuring outcomes.* Restructuring outcomes refers to the broad array of desired outcomes of a restructuring effort. To be comprehensive, a school restructuring design seeks not only to achieve school-wide impact, but also to improve several school outcomes in addition to student achievement. CSR legislation requires participating schools to implement models that address teacher support and involvement in the program, as well as parental involvement (U.S. Department of Education, 2007). Many of the models also seek to enhance the sense of community within the school, increase teacher collaboration, increase student independence in the learning process, enhance community support for the school, and address the educational needs of special needs students (Borman et al., 2004; Datnow & Stringfield, 2000). Although improved student achievement is best measured in other ways, teacher perceptual data are often collected to measure this and other restructuring outcomes (e.g., Berends et al., 2002).

### ***Overview and Purpose***

Failure to account for instructional practices and other mediating variables results in a “black box” model of studying school reform efforts (Olsen & Kirtman, 2002). While many studies have included classroom observation and teacher perception data as either outcomes or mediating variables in the evaluation of CSR models (e.g., Faddis et al., 2000), common instruments have not been consistently employed across studies. Many researchers relied on qualitative methods when observing classrooms in CSR schools (e.g., Bodilly, 1996; Datnow, 2000a, 2000b). Although collecting in-depth

qualitative observational data offers many advantages, a distinct disadvantage is the difficulty of comparing results across classrooms and schools, or directly linking restructuring changes to outcomes. Reliable and valid measures of school restructuring processes that can be employed across multiple designs would provide important tools for both the formative evaluation of ongoing restructuring efforts and the advancement of knowledge about school restructuring.

Thus, the purposes of the current study were to: (a) assess the construct validity of the *Comprehensive School Reform Teacher Questionnaire* (CSRTQ; Ross et al., 1999) as a measure of five broad constructs underlying teacher perceptions of school restructuring (*external support, internal capacity, focus, pedagogical change, and restructuring outcomes*); (b) estimate the school-level reliability of the *School Observation Measure* (SOM; Ross et al., 1998) as a measure of implementation of various pedagogical strategies; and (c) validate measures of pedagogical practice by estimating relationships between aggregate teacher perceptions and aggregate observation measures of instructional practices at the school level. To these ends, the following research questions were addressed:

1. How well do data collected via the CSRTQ fit the five-factor model of teacher perceptions of school restructuring?
2. Do aggregate (i.e., average ratings across one year of observations) SOM ratings reliably distinguish among schools with respect to instructional practices?
3. Given an affirmative finding for research question 2, do aggregate teacher perceptions of instructional reforms as measured by the CSRTQ *Pedagogical Change* subscale predict aggregate SOM item ratings as expected, viz.:
  - a. positive, significant slopes for SOM items denoting “reformed” instructional practices, such as project-based learning, cooperative learning, teacher as coach/facilitator, performance assessment, using the computer as a learning tool, and frequency of student discussion of instructional topics;
  - b. negative, significant slopes for SOM items denoting “traditional” instructional practices, such as independent seatwork or direct instruction.

## Methods

### *Participants*

Participants included 2,511 teachers in 83 schools that received Comprehensive School Reform (CSR) program grants beginning in the 1999-2000 school year. Participating schools were all those that contracted with AEL (formerly the Appalachian Education Laboratory) to provide evaluation services for their CSR programs. The schools were located in predominately rural school districts in Kentucky, Tennessee, Virginia, and West Virginia. Of these, 58 were elementary schools, 17 were middle or junior high schools, and eight were high schools. More than two-thirds (68.7%,  $n = 57$ ) were eligible Title 1 schoolwide project schools.

### *Data Collection Procedures and Instrumentation*

The *Comprehensive School Restructuring Teacher Questionnaire* (CSRTQ; Ross & Alberg, 1999) was administered to teachers at faculty meetings during the spring of the year following receipt of the CSR grant. The CSRTQ is comprised of 28 Likert-type items that solicit teachers' perceptions of the degree to which (a) implementation of the CSR reform has affected various outcomes in the school (e.g., student achievement, teacher collaboration); (b) the level of external support for implementation of the reform (quality of professional development, guidance from external facilitator); (c) school restructuring capacity (e.g., availability of resources, sufficient planning time); (d) the degree to which teachers are utilizing non-traditional pedagogical strategies advocated by many restructuring designs, such as cooperative learning and interdisciplinary projects; and (e) the degree to which the school has focused efforts and resources on implementing the comprehensive school reform strategy. On each item, teachers are asked to indicate the degree to which they agree with each statement on a 5-point Likert-type scale (1 = Strongly Disagree, 2 = Disagree, 3 = Not Sure, 4 = Agree, 5 = Strongly Agree). Items of the CSRTQ are listed by construct in the Table. Construct-item correspondence was determined *a priori*. Response rates ranged from 92% to 100% at individual schools, with a mean response rate of 96%. Approximately 2.2% of all data elements on the CSRTQ were missing due to a random pattern of nonresponse. Multiple imputation procedures using the EM-algorithm were employed to impute missing values (Schafer, 1997). The EM-algorithm converged in four iterations.

## Table

*Prediction of School Observation Measure Means from Teacher Self-reports of Pedagogical Change*

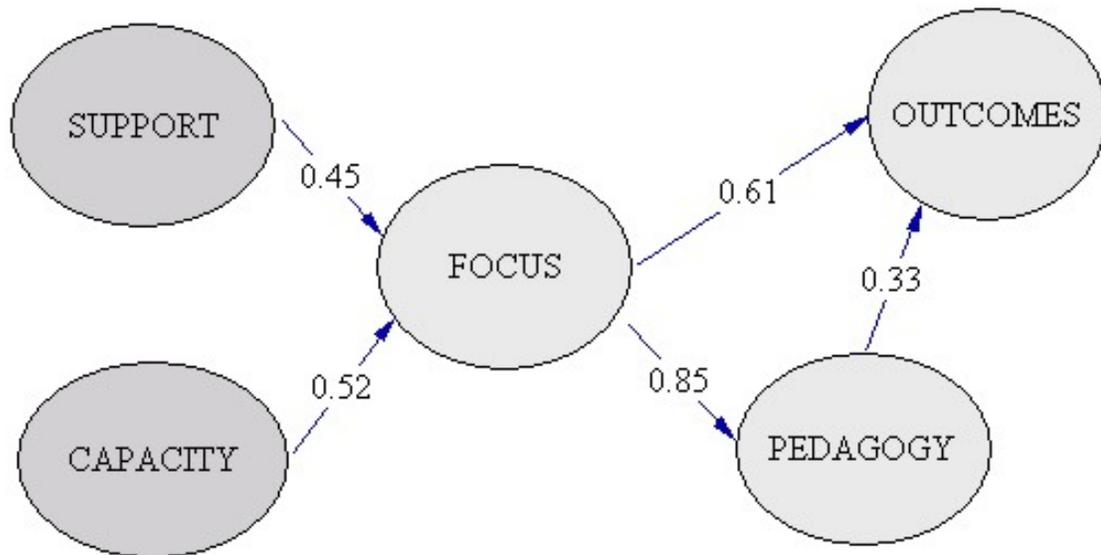
SOM Variable	Reliability <sup>1</sup>	$\gamma_p^2$	$t^3$	$p$
Direct instruction with whole class	0.68	-0.03	-0.22	0.83
Team teaching	0.86	0.18	1.32	0.19
Cooperative learning	0.79	0.82	5.67	<0.01
Individual tutoring	0.75	0.22	1.74	0.08
Ability groups	0.87	0.39	1.63	0.10
Multi-age grouping	0.92	0.77	3.16	<0.01
Work centers in use	0.81	0.69	4.29	<0.01
Instructional feedback	0.79	0.44	2.16	0.03
Integration of subject areas	0.80	0.34	2.41	0.02
Project-based learning	0.71	0.01	0.10	0.90
Higher-level questioning strategies	0.84	0.32	1.85	0.06
Teacher as coach or facilitator	0.80	0.74	4.30	<0.01
Parent or community involvement in learning activities	0.81	0.09	0.67	0.51
Independent seatwork	0.74	-0.31	-2.01	0.04
Experiential, hands-on learning	0.70	0.27	2.11	0.03
Systematic individual instruction	0.80	0.23	1.94	0.05
Sustained writing/composition	0.71	0.24	2.42	0.02
Sustained reading	0.77	0.08	0.77	0.44
Independent inquiry or research	0.75	0.10	0.90	0.37
Student discussion	0.80	0.28	1.94	0.05
Computer for instructional delivery	0.75	0.18	1.59	0.11
Computer as a learning tool	0.83	0.06	0.43	0.67
Performance assessment strategies	0.77	0.27	2.48	0.01
Student self-assessment	0.79	0.19	1.70	0.09
Academically focused class time	0.72	0.10	1.15	0.25
Student engagement/interest	0.73	0.17	1.93	0.05

Note.  $df_{(Level 2)} = 87$ . <sup>1</sup>Reliability of the school-level means. <sup>2</sup>Slope associated with the regression of variable on mean "Pedagogical Change" score from CSRTQ. <sup>3</sup>Computed with robust standard errors.

The *School Observation Measure* (SOM; Ross et al., 1998) was administered by trained observers on several occasions throughout the year at each participating school site. After conducting ten 15-minute observations of classroom instruction at a school site, observers rated the extent to which different common and alternative instructional practices were used in the school using a 5-point, Likert-type scale (1 = None or not observed, 2 = Rarely observed, 3 = Occasionally observed, 4 = Frequently observed, 5 = Extensively observed). In addition, the instrument solicits summary observers participated in extensive training. All received a manual providing definitions of terms, examples and explanations of target strategies, and a description of procedures for completing the instrument. After reviewing the manual and receiving instruction in a group session, each observer participated in practice exercises to ensure that his or her responses were comparable with those of experienced observers. In one reliability study (Lewis et al., 1999), pairs of trained observers selected the identical overall response on the five-category rubric on 67% of the items and were within one category on 95% of the items. In a second reliability study using generalizability theory, Sterbinsky and Ross (2003) found reliability at the .74 level for five SOMs conducted at a school. Reliability increased to .82 with eight SOMs and to .85 with ten SOMs conducted at a school. Accordingly, SOM observations at the schools participating in the present study were completed between eight and ten times at each school throughout the course of the school year (approximately one observation per month).

### ***CSRTQ Structural Model Specification***

A structural equation model comprised of observed (item responses) and latent (constructs) variables was estimated using the LISREL 8 program (Jöreskog, & Sörbom, 1996). Items assumed to vary as a function of variation in their respective latent construct were restricted to loading only on that construct. The model also specified *a priori* relationships among the latent constructs: *internal capacity* and *external support* were assumed to be uncorrelated; *internal capacity* and *external support* were assumed to have a direct effect on *internal focus*; *internal focus* was assumed to have direct effects on both *restructuring outcomes* and *pedagogical change*; and *pedagogical change* was assumed to have a direct effect on *restructuring outcomes* (see Figure 1). Because the item responses were on an ordinal scale, the asymptotic



*Figure 1.* Standardized path coefficients between latent factors underlying teachers' perceptions of comprehensive school reform.

covariance matrix was analyzed using maximum likelihood estimation method (Jöreskog & Sörbom, 1999).

### ***Estimation of School-Level SOM Reliability and Prediction of Aggregate SOM Scores from Aggregate CSRTQ Scores***

A two-level hierarchical linear model (HLM; Bryk & Raudenbush, 1992) was estimated for each SOM item. The level one (within-schools) model estimated the mean observer rating for each item over eight to ten observations:

$$Y_{ij} = \beta_{0j} + r_{ij};$$

where  $Y_{ij}$  is an individual rating on a SOM item at time  $i$  in school  $j$ ,  $\beta_{0j}$  is the average rating on a SOM item over all time periods in school  $j$ , and  $r_{ij}$  is a random error component. The level one model produces a single mean SOM rating for each school over the course of a year, as well as an average reliability estimate for the  $\beta_{0j}$  (Bryk & Raudenbush, 1992). The level two (between schools) model predicted mean SOM ratings using aggregate teacher perceptions of pedagogical change in the school during the same year, measured as the grand mean on the *Pedagogical Change* (PC) subscale of the CSRTQ:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} (\text{PC}) + u_{0j};$$

where  $\beta_{0j}$  is the estimated mean SOM item observer rating obtained in the

Level 1 model,  $\gamma_{00}$  is the estimated grand mean SOM rating for all schools,  $\gamma_{01}$  is the prediction slope of the  $\beta_{0j}$  versus aggregate PC subscale scores, and  $u_{0j}$  is a residual error term.

## Results

### *Structural Equation Model Fit*

The adjusted goodness-of-fit index (AGFI) value was 0.98, indicating a strong fit of the model to the observed data. The root mean square error of approximation was equal to 0.053, also indicating a close fit of the model to the data at  $p < .01$ . The parsimony normed fit index (PNFI), which adjusts the fit index based on degrees of freedom, was 0.87. Values obtained on these indices indicated that the *a priori* model provided a close and parsimonious fit to the observed data.

### *Measurement Equations*

The appendix provides an overview of the measurement equations relating each item to its respective construct. The equations indicate how well variation in the items is predicted by levels of the underlying construct. Standardized regression coefficients ( $\lambda$ ) of the items on the latent constructs ranged from 0.21 to 0.75, with a median value of 0.28. All  $\lambda$ s were significant at  $p < .001$ .  $R^2$  for individual items ranged from 0.18 to 0.66, with a median value of 0.39. Generally,  $\lambda$ s were larger for *support* and *capacity* items, suggesting that teacher responses may be more sensitive to variations in these constructs. Figure 1 portrays the standardized structural equation model among the latent constructs. Perceptions of *focus* were found to be nearly equally sensitive to perceived *support* ( $\gamma = 0.45$ ) and *capacity* ( $\gamma = 0.52$ ), which together accounted for 84% of the variance in *focus*. *Focus* had strong direct effects on both *pedagogy* ( $\gamma = 0.85$ ) and *outcomes* ( $\gamma = 0.61$ ). Together, *focus* and *pedagogy* ( $\gamma = 0.33$ ) explained 84% of the variance in *outcomes*. At the individual teacher level, external support and internal capacity had relatively strong and nearly equal effects on internal focus, while focus and pedagogical change had strong direct effects on restructuring outcomes.

### ***Reliability of Aggregate SOM Ratings***

Average reliability estimates for SOM item ratings ranged from 0.68 to 0.92, with a median value of 0.79 (see Table). “Direct instruction with the entire class” and “Experiential, hands-on learning” had the lowest reliability estimates (0.68 and 0.70, respectively), whereas “Multi-age grouping” and “Ability grouping” had the highest reliability estimates (0.92 and 0.87, respectively). For every item, more than two-thirds of the variance lay between schools, indicating that all aggregate SOM item ratings suitably distinguished between schools.

### ***Prediction of SOM Ratings from Aggregate CSRTQ Pedagogical Change Scores***

A total of 26 SOM/PC slopes were tested for significance. The most conservative approach to testing for significance would be to use a Bonferroni adjustment, setting  $\alpha$  equal to  $0.05/26 = 0.0019$ . Using the Bonferroni approach, substantial and statistically significant relationships were observed between PC scores and four SOM items: *cooperative learning*,  $\gamma = 0.82$ ,  $t(87) = 5.67$ ,  $p < .001$ ; *multi-age grouping*,  $\gamma = 0.92$ ,  $t(87) = 3.16$ ,  $p < .001$ ; *work centers in use*,  $\gamma = 0.69$ ,  $t(87) = 4.29$ ,  $p < .001$ ; and *teacher as coach or facilitator*,  $\gamma = 0.80$ ,  $t(87) = 4.30$ ,  $p < .001$ ; see Table 2. Using a less conservative criterion (i.e.,  $\alpha = .05$  for each comparison), significant, positive relationships were observed between aggregate PC scores and several additional SOM items: *instructional feedback*,  $\gamma = 0.44$ ,  $t(87) = 2.16$ ,  $p = .03$ ; *integration of subject areas*,  $\gamma = 0.34$ ,  $t(87) = 2.41$ ,  $p = .02$ ; *experiential, hands-on learning*,  $\gamma = 0.27$ ,  $t(87) = 2.11$ ,  $p = .03$ ; *systematic individual instruction*,  $\gamma = 0.23$ ,  $t(87) = 1.94$ ,  $p = .05$ ; *sustained writing or composition*,  $\gamma = 0.24$ ,  $t(87) = 2.42$ ,  $p = .02$ ; *student discussion*,  $\gamma = 0.28$ ,  $t(87) = 1.94$ ,  $p = .05$ ; *performance assessment strategies*,  $\gamma = 0.27$ ,  $t(87) = 2.48$ ,  $p = .01$ ; and *student engagement/interest*,  $\gamma = 0.17$ ,  $t(87) = 1.93$ ,  $p = .05$ ; see Appendix. A statistically significant, negative relationship between aggregate PC scores and SOM item ratings was observed for *independent seatwork*,  $\gamma = -0.31$ ,  $t(87) = -2.01$ ,  $p = .04$ . Overall, aggregate teacher perceptions of pedagogical change were good predictors of observed teaching practice—PC scores were positively related to the extent to which “reformed” teaching practices were observed, and negatively related to the extent to which “traditional” teaching practices were observed.

## Discussion

In response to the first research question, we found the data collected via the CSRTQ supported the five-factor model of teacher perceptions of school restructuring. This study provides construct validation evidence for the CSRTQ, an instrument designed to measure global constructs underlying teacher perceptions of comprehensive school reform. Although the CSR legislation (U.S. Department of Education, 2007) encourages schools to adopt “proven” designs, the model formulated and tested in this study suggests that the professional development, materials, and technical assistance provided by a design team are only likely to be effective if the school has sufficient capacity and internal focus to implement the reform. Several other researchers have also reported that capacity and internal focus are crucial predictors of the successful school reform (Bodilly, 1996; Borman et al., 2004; Datnow & Stringfield, 2000; Desimone, 2002; Rowan et al., 2004). These findings lend further support to earlier findings linking teacher perceptions about capacity and focus to positive outcomes in schools implementing comprehensive school reform (Bol et al., 1998; Ross et al., 1997; 2000; Stringfield & Datnow, 1998; Supovitz & May, 2004). The five-factor model instantiated in the CSRTQ can provide a framework and a tool for research on the interaction of school and design characteristics in creating focus, changing instructional practices, and improving school outcomes.

In response to our second research question, we found that SOM ratings reliably distinguished between schools with respect to instructional practices. Specifically, the aggregate teacher perceptions of pedagogical change as measured on the CSRTQ are reliable and valid indicators of school-level instructional practice, in that they predict scores obtained through extensive high-inference observations of instructional practice in the school. These results not only support the measurement properties of the CSRTQ but also suggest that the SOM observations are a viable tool for linking pedagogical practice, teacher perceptions, and student outcomes in more sophisticated model for understanding of the impact of CSR initiatives. Earlier studies have also shown that variations in pedagogical practices distinguish schools and help to explain outcomes in schools undergoing reform (Hamilton et al., 2003; Ross et al., 2000).

Not only did teacher perceptions as measured by *Pedagogical Change* scores reliably predict SOM ratings, the results were well aligned with

predictions based on effective pedagogical practices. In schools with relatively high aggregate PC scores, reformed practices such as cooperative learning, teacher acting as coach or facilitator, experiential learning, and subject area integration were observed more frequently, while independent seatwork was observed less frequently. Higher PC scores were also related to more frequent observation of sustained writing, student discussion, and student engagement. These findings suggest that schools in which teachers reported greater levels of pedagogical change as a consequence of implementing CSR initiatives, which generally call for more student-centered instructional practices (Bodilly, 1996; Ross, et al., 2000; Stringfield et al., 1996), were in fact implementing these instructional practices.

The present findings further illustrate that teachers who perceive a clear focus on reform within the school are likely to report the greatest changes in instructional practice and restructuring outcomes. This suggests that effective planning, teacher buy-in, program integration, regular review of progress, and coordination of external resources are instrumental in producing changes in pedagogy. Teacher perceptions of both internal capacity (time, materials, staffing, equipment) and external support (professional development, technical assistance) seem strongly related to the degree to which teachers report development of a reform-oriented focus in their schools. The literature is replete with both theory and findings that attest to the importance of both internal focus (e.g., Berends et al., 2002; Metz, 1988; Ross et al., 2000) and external support (e.g., Bol et al., 1998; Desimone, 2002; Hopkins et al., 1994), providing the basis for components of CSR program guidelines (US Department of Education, 2002). When implementing these programs, schools, districts, and redesign teams should consider not only the match of design characteristics to school goals and educational philosophy, but to the internal capacity of the school to engage in the reform effort.

Despite these encouraging findings, there are some limitations that bear note. Because the teacher sample was drawn largely from elementary, rural, Title 1 schools, the results may not generalize to other contexts. Further, these schools and districts proactively sought external resources to implement restructuring models, and thus may differ in important ways from schools wherein change is externally mandated. Although the *a priori* model proved to fit the data quite well, it is possible that other specifications could provide an equally good fit.

In conclusion, our findings indicate that teachers' perceptions of change are valid predictors of the impact of comprehensive school reform models on instructional practice, and suggest that the CSRTQ and SOM instruments can reliably distinguish between schools with respect to implementation of student-centered instructional practices. Together, the two instruments can serve as practical, low-cost formative assessment tools for schools undertaking implementation of school-wide reforms, and can also help advance knowledge about restructuring implementation and outcomes by shining light in the "black box" of educational reform (Olsen & Kirtman, 2002). Future research using the CSRTQ and SOM might focus on the degree to which they are predictive of improvement in student achievement, and on the interaction between local context and restructuring design characteristics.

## References

- Berends, M., Bodilly, S., & Kirby, S. N. (2002). Facing the challenges of whole-school reform: The experience of New American Schools. In J. Murphy & A. Datnow (Eds.). *Leadership lessons from comprehensive school reforms* (pp. 109-131). Thousand Oaks, CA: Corwin Press.
- Berman, P., & McLaughlin, M. W. (1977). *Federal programs supporting educational change. Vol. VI. Implementing and sustaining innovations*. Santa Monica, CA: RAND.
- Bodilly, S. (1996). *Lessons from New American School Development Corporation's demonstration phase*. Santa Monica, CA: Rand.
- Bodilly, S. (2001). *New American Schools' concept of break the mold designs: How designs evolved over time and why*. Santa Monica, CA: Rand.
- Bol, L., Nunnery, J. A., Lowther, D. L., Dietrich, A. P., Pace, J., Anderson, R. S., et al. (1998). Inside-in and outside-in support for restructuring: The effects of internal and external support on change in the New American Schools. *Education and Urban Society, 30*, 358-384.
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125-230.

- Borman, K. M., Carter, K., Aladjem, D. K., & LeFloch, K. C. (2004). Challenges for the future of comprehensive school reform. In C.T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 53-108). Washington, DC: The National Clearinghouse for Comprehensive School Reform.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 119-142.
- 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*, 30(3), 385-408.
- Crandall, D. P., & Loucks, S. F. (1983). *A roadmap for school improvement: An executive summary of the study of dissemination efforts supporting school improvement* (Vol. 10). Andover, MA: NETWORK.
- Datnow, A. (2000a). Implementing an externally developed school restructuring design: Enablers, constraints, and tensions. *Teaching and Change*, 7, 147-171.
- Datnow, A. (2000b). Power and politics in the adoption of school reform models. *Educational Evaluation and Policy Analysis*, 22, 357-374.
- Datnow, A., & Stringfield, S.C. (2000). Working together for reliable school reform. *Journal of Education for Students Placed At Risk*, 5, 183-204.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72, 433-479.
- Faddis, B. J., Beam, M., Hahn, K. J., Willardson, M., Sipe, D., & Ahrens-Gray, P. (2000). *The implementation of the Comprehensive School Reform Demonstration Project: The work of 40 schools in seven Midwest states*. Naperville, IL: North Central Regional Laboratory.
- Fuhrman, S. H., Clune, W., & Elmore, R. (1988). Research on educational reform: Lessons on the implementation of policy. *Teachers College Record*, 90(2), 237-257.
- Fullan, M. (1991). *The new meaning of educational change*. London: Cassell.

- Glennan, Jr., T. K. (1998). *New American Schools after six years*. Santa Monica, CA: RAND.
- Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., Klein, S. P., Robyn, A., & Bugliari, D. (2003). Studying large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis, 25*, 1-29.
- Havelock, R. G. (1971). *Planning for innovation through utilization and dissemination of knowledge*. Ann Arbor: Center for Research on Utilization of Scientific Knowledge, University of Michigan.
- Hopkins, D., Ainscow, M., & West, M. (1994). *School improvement in an era of change*. London: Cassell.
- Jöreskog, K. G., & Sörbom, D. (1999). *LISREL 8 user's reference guide*. Chicago: Scientific Software International.
- Lewis, E. M., Ross, S. M., & Alberg, M. J. (1999). *School Observation Measure: Reliability analysis*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.
- McLaughlin, M. W. (1990). The RAND Change Agent study revisited: Macro perspectives and micro realities. *Educational Researcher, 19*, 11-16.
- McLaughlin, M., & Talbert, J. (2003). *Reforming districts: How districts support school reform*. Seattle, WA: Center for Study of Teaching and Policy.
- Metz, M. H. (1988). Some missing elements in the educational reform movement. *Educational Administration Quarterly, 24*(4), 446-460.
- Nunnery, J. A., Bol, L., Dietrich, A., Rich, L., Kelly, S., Hacker, D., & Sterbin, A. (1997). Teachers' initial reactions to their pre-implementation preparation and early restructuring experiences. *School Effectiveness and School Improvement, 81*, 72-94.
- Nunnery, J. A. (1998). Reform ideology and the locus of development problem in educational restructuring. *Education and Urban Society, 30*, 277-295.
- Olsen, B., & Kirtman, L. (2002). Teacher as mediator of school reform: An examination of teacher practice in 36 California restructuring schools. *Teachers College Record, 104*, 301-324.
- Puma, M. J., & Drury, D. W. (2000). *Exploring new directions: Title I in the Year 2000*. Alexandria, VA: National School Boards Association.

- Ross, S. M., & Alberg, M. J. (1999). *The Comprehensive School Reform Teacher Questionnaire*. Memphis, TN: Center for Research in Educational Policy.
- Ross, S. M., Alberg, M. J., & Nunnery, J. A. (1999). Selection and evaluation of locally-developed versus externally-developed schoolwide programs. In G. Orfield and E.G. DeBray (Eds.), *Hard work for good schools: Facts not fads in Title I reform* (pp. 147-158). Cambridge, MA: The Civil Rights Project, Harvard University.
- Ross, S. M., Alberg, M., Smith, L., Anderson, R., Bol, L., Dietrich, A., Lowther, D., & Phillipsen, L. (2000). Using whole-school restructuring designs to improve educational outcomes: The Memphis story at Year 3. *Teaching and Change*, 7, 111-126.
- Ross, S. M., & Gil, L. (2004). The past and future of comprehensive school reform: Perspectives from a researcher and practitioner. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 151-174). Washington, DC: The National Clearinghouse for Comprehensive School Reform.
- Ross, S. M., Sanders, W., Wright, S., Stringfield, S., Wang, W., & Alberg, M. (2001). Two- and three-year achievement results from the Memphis restructuring initiative. *School Effectiveness and School Improvement*, 12, 323-46.
- Ross, S. M., Smith, L. J., & Alberg M. J. (1998). *School Observation Measure*. Memphis, TN: Center for Research in Educational Policy, The University of Memphis.
- Ross, S. M., Smith, L. J., Alberg, M., & Lowther, D. L. (2004). Using classroom observation as a research and formative evaluation tool in educational reform: The School Observation Measure. In H. Waxman (Ed.), *Observational research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity* (pp. 144-173). Cambridge, MA: Cambridge University Press.
- Ross, S. M., Troutman, A., Horgan, D., Maxwell, S., Laitinen, R., & Lowther, D. (1997). The success of schools in implementing eight restructuring designs: A synthesis of first-year evaluation outcomes. *School Effectiveness and School Improvement*, 8(1), 95-124.

- Rowan, B., Camburn, E., & Barnes, C. (2004). Benefiting for comprehensive school reform: A review of research on CSR implementation. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 1-52). Washington, DC: The National Clearinghouse for Comprehensive School Reform.
- Schafer, J. L. (1997). Analysis of incomplete multivariate data. *Mono-graphs on statistics and applied probability*, 72, Chapman and Hall/CRC.
- Slavin, R. E. (1999). How Title I can become the engine of reform in America's schools. In G. Orfield & E. H. DeBray (Eds.), *Hard work for good schools: Facts, not fads, in Title I reform* (pp. 86-102). Cambridge, MA: Harvard University, Civil Rights Project.
- Smith, L. J., Maxwell, S., Lowther, D., Hacker, D., Bol, L., & Nunnery, 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. J., Ross, S. M., McNelis, M., Squires, M., Wasson, R., Maxwell, S., Weddle, K., Nath, L., Grehan, A., & Buggey, T. (1998). The Memphis restructuring initiative: Analysis of activities and outcomes that affect implementation success. *Education and Urban Society*, 30, 296-325.
- Sterbinsky, A., & Ross, S. M. (2003). *School Observation Measure reliability study*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.
- Stringfield, S. C., & Datnow, A. (1998). Scaling up school restructuring designs in urban schools. *Education and Urban Society*, 30(3), 269-276.
- Stringfield, S. C., Ross, S. M., & Smith, L. J. (1996). *Bold plans for school restructuring: The New American Schools designs*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stringfield, S. C., Datnow, A., Ross, S. M., & Snively, F. (1998). Scaling up school restructuring in multicultural, multilingual contexts: Early observations from Sunland County. *Education and Urban Society*, 30, 326-357.
- Stringfield, S. C., Waxman, H.C., & Padron, Y. N. (2000). To be of lasting quality, school reform must include change in teaching. *Teaching and Change*, 7, 107-110.

- Supovitz, J. A., & May, H. (2004). A study of the links between implementation and effectiveness of America's Choice comprehensive school reform design. *Journal of Education for Students Placed at Risk*, 9, 389-420.
- U.S. Department of Education. (2007). Overview of the Comprehensive School Reform Program. Retrieved April 10, 2008 from <http://www.ed.gov/programs/compreform/2pager.html>

## Appendix

### *Comprehensive School Restructuring Questionnaire: Confirmatory Factor Analysis Measurement Equations*

Factor/ Items	$\lambda$	$t(\lambda)$	$\delta$	$R^2$
<i>Support</i>				
Professional development provided by external trainers, model developers, or designers has been valuable.	0.74	43.19	0.36	0.60
Guidance and support provided by our school's external facilitator, support team, or other state-identified resource personnel have helped our school implement its program.	0.72	42.55	0.37	0.59
I have received adequate professional development for implementation.	0.75	38.60	0.51	0.52
I have a thorough understanding of this school's CSR program.	0.69	34.30	0.63	0.43
My school receives effective assistance from external partners.	0.40	21.18	0.70	0.18
<i>Capacity</i>				
Teachers are given sufficient planning time to implement our model.	0.65	26.88	0.94	0.31
Materials needed to implement our CSR program are readily available.	0.74	32.05	0.82	0.29
Our school has sufficient faculty and staff to implement this program.	0.58	26.09	0.82	0.29

Factor/ Items	$\lambda$	$f(\lambda)$	$\delta$	$R^2$
<i>Capacity (continued)</i>				
Technological resources have become more available.	0.54	26.26	0.77	0.28
<i>Focus</i>				
Teachers in this school are generally supportive of our CSR program.	0.26	22.07	0.38	0.52
Our school has a plan for evaluating all components of our CSR program.	0.21	21.09	0.36	0.42
The elements of our CSR program are effectively integrated to help us meet school improvement goals.	0.26	22.96	0.48	0.37
As a school staff, we regularly review implementation and outcome benchmarks to evaluate our progress.	0.22	20.48	0.48	0.37
I am satisfied that federal, state, local, and private resources are being coordinated to support our CSR program.	0.22	20.68	0.48	0.37
<i>Pedagogy</i>				
Our CSR program has changed classroom learning activities a great deal.	0.37	24.69	0.35	0.59

Factor/ Items	$\lambda$	$t(\lambda)$	$\delta$	$R^2$
<i>Pedagogy (continued)</i>				
Because of our CSR program, I use textbooks, workbooks, and worksheets less than I used to for basic skills or content area instruction.	0.33	23.00	0.56	0.42
Students are using technology more effectively because of our CSR program.	0.27	21.31	0.68	0.29
Children in my class spend at least two hours per school day in interdisciplinary or project-based work.	0.26	20.17	0.67	0.28
Students in my class spend much of their time working in cooperative learning teams.	0.24	18.27	0.77	0.21
<i>Outcomes</i>				
Children in this school are more enthusiastic about learning than they were before we became a CSR school.	0.31	29.09	0.31	0.66
Student achievement has been positively impacted by CSR.	0.29	28.69	0.29	0.63
Because of CSR, interactions between teachers and students are more positive.	0.27	28.55	0.28	0.62
Students have higher standards for their own work because of our school's program.	0.29	28.41	0.34	0.61

Factor/ Items	$\lambda$	$t(\lambda)$	$\delta$	$R^2$
<i>Outcomes (continued)</i>				
Teachers are more involved in decision making at this school than before CSR.	0.27	25.50	0.59	0.43
Community support for our school has increased since CSR has been implemented.	0.24	25.35	0.49	0.42
Because of CSR, parents are more involved in the educational program in this school.	0.24	24.84	0.55	0.39
Teachers in this school spend more time working together to develop curriculum and plan instruction.	0.26	24.72	0.64	0.39
Our program addresses the requirements of children with special needs.	0.23	23.28	0.68	0.33

*Note.* All  $t$ -values significant at  $p < .001$ ;  $\lambda$  = standardized regression coefficient of the item on the latent factor ( $\xi_j$ );  $\delta$  = error variance.