Journal of School Psychology 51 (2013) 437–453

Contents lists available at SciVerse ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

The influence of fidelity of implementation on teacher–student interaction quality in the context of a randomized controlled trial of the *Responsive Classroom* approach ☆,☆☆

Tashia Abry [a,*], Sara E. Rimm-Kaufman [b], Ross A. Larsen [b], Alexis J. Brewer [b]

[a] *Arizona State University, USA*
[b] *University of Virginia, USA*

ARTICLE INFO

Article history:
Received 8 September 2011
Received in revised form 15 March 2013
Accepted 18 March 2013

Keywords:
Fidelity of implementation
Teacher–student interaction
Social and emotional learning
Structural equation modeling
Responsive Classroom approach
Teacher quality

ABSTRACT

This study examined the direct and indirect effects between training in the *Responsive Classroom*® (*RC*) approach, teachers' uptake of *RC* practices, and teacher–student interaction quality, using a structural equation modeling framework. A total of 24 schools were randomly assigned to experimental or control conditions. Third- and fourth-grade teachers in treatment schools ($n = 132$) received training in the *RC* approach, whereas teachers in control schools ($n = 107$) continued "business as usual." Observers rated teachers' fidelity of implementation (FOI) of *RC* practices 5 times throughout the year using the Classroom Practices Observation Measure. In addition, teachers completed self-report measures of FOI, the Classroom Practices Teacher Survey and Classroom Practices Frequency Survey, at the end of the school year. Teacher–student interactions were rated during classroom observations using the Classroom Assessment Scoring System. Controlling for teachers' grade level and teacher–student interaction quality at pretest, *RC* training was expected to predict posttest teacher–student interaction quality directly and indirectly through FOI. Results supported only a significant indirect effect, $\beta = 0.85$, $p = .002$. Specifically, *RC* teachers had higher levels of FOI of *RC* practices, $\beta = 1.62$, $p < .001$, $R^2 = .69$. In turn, FOI related to greater improvement in teacher–student interaction quality, $\beta = 0.52$, $p = .001$, $R^2 = .32$. Discussion highlights factors contributing to variability in FOI and school administrators roles in supporting FOI.

1. Introduction

School-age children spend nearly half of their waking hours in school. For this reason, classrooms have become a critically important setting in which to intervene to improve the social and academic performance of American youth (Greenberg, 2010). Social and emotional learning (SEL) programs offer specialized professional development to teachers to improve the classroom climate, students' social skills, and ultimately academic achievement (Zins, Weissberg, Wang & Walberg, 2004). Indeed, recent meta-analytic work on SEL interventions has shown that students across elementary and secondary grades demonstrate greater

☆ The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through grant R305A070063 to the second author and by the Institute of Education Sciences, U.S. Department of Education, through grants R305B090002 and R305B100009 to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.
☆☆ This paper was presented in an earlier form at the 2012 annual meeting of the American Education Research Association. An earlier version of this paper is included as a component of the first author's doctoral dissertation. The authors wish to thank Julie Thomas, Michelle Ko, Jennifer Williams, and Abby Moncrief as well as collaborating district administrators and teachers.
* Corresponding author at: T. Denny Sanford School of Social and Family Dynamics, Arizona State University, PO Box 873701, Tempe, AZ, 85287, USA. Tel.: +1 480 727 9205.
E-mail address: tabry@asu.edu (T. Abry).
ACTION EDITOR: Craig Albers.

0022-4405/$ – see front matter © 2013 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jsp.2013.03.001

social–emotional outcomes and greater academic outcomes compared to control students with effect sizes (i.e., Cohen's *d*) of 0.57 and 0.27, respectively (Durlak, Weissberg, Dymnicki, Taylor & Schellinger, 2011). Increasingly, schools are adopting such programs in response to local, state, and federal policies (Durlak et al., 2011).

Demonstrating the link between SEL interventions and targeted outcomes is crucial to establish program effectiveness and justify the continued inclusion of such interventions in school programming (Kress & Elias, 2006). With a growing evidence base substantiating the direct link between SEL programs and positive youth development, attention has turned to explaining the mechanisms of change underlying such interventions (Karachi, Abbott, Catalano, Haggerty & Fleming, 1999). Unpacking such mechanisms requires testing hypothesized mediators specified in the logic model of the intervention (Nelson, Cordray, Hulleman, Darrow & Sommer, 2012; O'Donnell, 2008). However, analyses of child outcomes have remained the predominant focus in most studies of SEL programs (Domitrovich & Greenberg, 2000), with just a handful of exceptions (see Brown, Jones, LaRusso & Aber, 2010; Conduct Problems Prevention Research Group [CPPRG], 1999). As a result, surprisingly little is known about how the adoption of SEL programs impacts day-to-day interactions between teachers and students, one of the primary mechanisms presumed to explain gains in youth outcomes (Zins, Bloodworth, Weissberg & Walberg, 2007).

Examination of the effectiveness of a SEL intervention also requires attention to fidelity of implementation (Dane & Schneider, 1998). For example, when intervention effects are present, measures of fidelity of implementation assist researchers in identifying the core components of the program to replicate in future studies and in large-scale dissemination. In the absence of treatment effects, measures of fidelity of implementation aid researchers in determining whether the null results reflect a failure of implementation or failure of the intervention. That is, implementers (e.g., teachers) may have simply failed to uptake the intervention. On the other hand, it is possible that implementers did adopt the intervention, but it did not have the anticipated impact. For these reasons, fidelity of implementation is an important explanatory variable to consider when examining targeted outcomes and fidelity measures are being included in program evaluations with increasing frequency (Century, Rudnick & Freeman, 2010).

Our goal in the present study was to understand the direct and indirect relations among teacher professional development in a SEL intervention, use of classroom practices associated with the professional development (i.e., fidelity of implementation), and teacher–student interaction quality in the context of a district-level mandate to adopt a behavior support program. We examined these associations within the *Responsive Classroom*® (*RC*) approach, a widely disseminated SEL intervention. Specifically, we assessed the direct effects of teachers' training in the *RC* approach on their use of *RC* practices and the quality of teacher–student interactions, as well as the indirect effect of *RC* training on interaction quality via teachers' implementation of *RC* practices. The goal was to elucidate the extent to which teachers adopted intervention practices subsequent to the professional development and whether the use of intervention practices related to improvements in quality of teacher–student interactions. Thereby, we attempted to disentangle the influence of training in an intervention from implementation of the intervention on improvements in classroom interactions.

1.1. The Responsive Classroom approach

The present study takes place in the context of a randomized controlled trial of the *RC* approach (Rimm-Kaufman et al., submitted for publication). The *RC* approach is designed to improve elementary children's social, emotional, self-regulatory, and academic development through the creation of a well-structured and supportive learning environment (Northeast Foundation for Children [NEFC], 2007). In its entirety, the *RC* approach is integrated throughout the elementary school; however, training for teachers and classroom implementation is the initial and sometimes sole focus, meaning that the *RC* approach can be examined as either a classroom- or school-level intervention. Since 1981, more than 90,000 kindergarten through sixth-grade teachers have been trained in the *RC* approach nationwide. Currently, approximately 6000 teachers attend workshops and trainings in the *RC* approach annually (NEFC, 2012a,b). Teachers receive training in principles and practices of the *RC* approach through a structured framework involving comprehensive weeklong training sessions, ongoing coaching support, instructional books, and newsletters. The training sessions are divided into two one-week institutes (*RC* 1 and *RC* 2), typically taken during two consecutive summers. *RC* 1 focuses on *RC* practices related most closely to the emotional and managerial climates of the classroom. These practices are expanded in *RC* 2 and there is an increased focus on practices related to instruction, problem solving, and parent involvement. NEFC staff provides individual coaching sessions throughout the school year subsequent to each training institute.

The *RC* approach encompasses seven guiding principles that highlight the importance of the school and familial context, social interaction and social skills, and process-oriented learning (NEFC, 2007). Ten classroom practices emanate from these guiding principles. Examples of practices include the following: (a) Morning Meeting — a daily circle-time in which students greet each other, share personal news, participate in a fun activity, and process an interactive message written by the teacher; (b) Rule Creation — a process by which teachers and students work together to distill a set of three to five global classroom rules that support student-generated academic and social goals for the year; (c) Interactive Modeling — a proactive approach to teaching rules and routine classroom behaviors in which teachers model a desired behavior, ask students to make observations about the demonstration, and allow students opportunities to practice and reflect on their practice; and (d) Academic Choice — a structured process by which teachers help students to plan, complete, and reflect on academic work of their own choosing (NEFC, 2007, 2012b). As a whole, the *RC* approach aims to create classroom conditions optimized for students' social, emotional, and self-regulatory skills with the ultimate goal of improving academic achievement.

1.1.1. Extant research on the RC approach

Quasi-experimental work examining the efficacy of the *RC* approach indicated its positive association with child-centered discipline and teaching priorities (Rimm-Kaufman, Storm, Sawyer, Pianta & LaParo, 2006), increased teacher collaboration (Sawyer &

Rimm-Kaufman, 2007), and enhanced teacher self-efficacy and improved attitudes toward teaching (Rimm-Kaufman & Sawyer, 2004). Additionally, quasi-experimental evidence has demonstrated the relation between the *RC* approach and student gains in reading (Cohen's $d = 0.16$ to 0.21) and math achievement (Cohen's $d = 0.16$ to 0.39) after at least two years of program exposure in third through fifth grades (Rimm-Kaufman, Fan, Chiu & You, 2007). The prevalence of the *RC* approach in elementary schools, combined with promising results from quasiexperimental studies, suggests that the approach is accessible to teachers and merits experimental evaluation.

1.2. The importance of classroom interactions

The quality of classroom processes such as the interactions between teachers and students has been linked to positive student developmental outcomes. Students benefit in a variety of ways when they feel welcome and respected, when classrooms are well-structured, and when teachers promote understanding by matching instruction to their individual and developmental needs. For example, elementary and middle school students who perceive their teachers as caring and supportive are reported to have better social competency and higher levels of motivation and interest in class activities (Howes, 2000; Wentzel, 1997, 2002), which link to later gains in academic achievement (Hughes, Luo, Kwok & Loyd, 2008; Klem & Connell, 2004).

Effective management, which is exemplified by proactive discipline strategies, maximized learning time, and effective use of techniques that encourage active student engagement (Emmer & Stough, 2001; Sugai & Horner, 2002), can help students entering school in kindergarten capitalize on the learning opportunities provided in the classroom by supporting students' self-control and behavioral engagement (Rimm-Kaufman, Curby, Grimm, Nathanson & Brock, 2009). As a result, academic growth has been shown with first-grade students to be steeper for students in well-structured classrooms compared to those with higher levels of chaos (Ponitz, Rimm-Kaufman, Brock & Nathanson, 2009).

Instructional interactions are also influential on students' social and academic skills. Active instruction in which teachers directly instruct students through lecture, discussion, and individual feedback (Brophy & Good, 1986) has been linked to growth in both reading and mathematics achievement in elementary grades and into middle school (Rowan, Correnti & Miller, 2002). Furthermore, instructional interactions marked by teachers' use of scaffolding, explicit feedback, and promotion of higher order thinking are positively associated with teacher reports of social competency, math, and literacy skills in kindergarten students (Pianta, LaParo, Payne, Cox & Bradley, 2002). Taken together, these findings suggest that interventions that can improve the interactions between teachers and students are also likely to enhance students' social and academic development.

1.3. Improving teacher–student interactions with SEL interventions

Recent work demonstrates teacher–student interaction quality is a malleable point of intervention (Pianta, Mashburn, Downer, Hamre & Justice, 2008c; Raver et al., 2008), raising questions about the extent to which SEL interventions improve such interactions. Conceptual models of SEL interventions identify improvements in classroom processes such as teacher and peer interactions as an outcome of program adoption (Zins et al., 2007), but surprisingly, little empirical work has examined this claim.

Two notable exceptions exist. The Fast Track Promoting Alternative Thinking Strategies (PATHS) curriculum is a classroom-based prevention program focused on identifying and communicating emotions, prosocial behavior, and problem solving and was found to improve observational ratings of classroom atmosphere (CPPRG, 1999). Specifically, four aspects of the overall classroom atmosphere, including students' level of interest, ability to follow rules, appropriate expression of feelings, and ability to stay on task, were found to be higher in a first-grade experimental group than in control classrooms (CPPRG, 1999). However, assignment to the PATHS intervention group did not predict other aspects of the classroom atmosphere including teachers' responsiveness to student needs and feelings, use of problem solving, and level of criticism. A second example is a recent experimental study of the effects of the 4Rs program, a SEL and literacy program, on the quality of teacher–student interactions in third grade (Brown et al., 2010). Findings showed that 4Rs teachers had higher-quality emotional and instructional interactions, marked by teacher sensitivity, positive classroom relationships, individualized feedback, and opportunities for student discussion compared to teachers in the control condition.

Research using quasi-experimental methods also supports the link between SEL programming and improvements in classroom level processes. For example, the School Development Program, a school-based reform model focused on improving interpersonal relationships and school climate as a vehicle toward advancing student achievement, has been shown to improve middle school students' perceptions of their school's social climate as early as the first year of program adoption (Cook, Murphy & Hunt, 2000). Similarly, analyses of the Child Development Project, a school-based intervention aimed at promoting students' social, ethical, and intellectual development, revealed that nearly all effects of the program were mediated through elementary students' increased sense of community (Solomon, Battistich, Watson, Schaps & Lewis, 2000). Together, these findings help establish the theoretical link between SEL interventions and improvements in the school and classroom environment. The present inquiry contributes to the small body of experimental work linking SEL programming to improvements in teacher–student classroom interactions and extends this work by assessing the role of program implementation as the vehicle for creating classroom change.

1.4. Fidelity of implementation

Randomized controlled trials provide a methodologically rigorous design for determining whether membership in a treatment group induces change in measured outcomes. However, such experimental trials are not without their limitations. As with any

intervention study, assignment to a treatment group does not ensure that the intervention is being delivered faithfully nor does it guarantee that aspects of the intervention are completely absent in the control group (Hulleman & Cordray, 2009). If not taken into consideration, either scenario decreases internal validity and can call into question the accuracy of conclusions made about program efficacy (Cook & Campbell, 1979; Cordray & Pion, 2006). The potential to misrepresent conclusions about program effectiveness underscores the importance of measuring the degree to which implementers actually use prescribed program components.

Fidelity of implementation (FOI), broadly defined, refers to the extent to which the core components of a program are delivered as intended by program developers (Century et al., 2010; Dusenbury, Brannigan, Falco & Hansen, 2003). FOI consists of multiple domains (Dane & Schneider, 1998) including *adherence* (how closely one follows prescribed program components) and *dosage or frequency* (the amount of the program implemented). For the purpose of this study, we operationalize FOI as the adherence to and frequency with which teachers use the practices introduced during the *RC* training in a way that is consistent with the guiding principles of the *RC* approach.

1.4.1. Professional development and fidelity of implementation

The professional development accompanying an intervention, including training and posttraining support, has been repeatedly associated with higher levels of FOI. For example, training attendance and perceived training quality are associated with higher levels of dosage and adherence to program components (Ransford, Greenberg, Domitrovich, Small & Jacobson, 2009; Ringwalt et al., 2003; Ross, Luepker, Nelson, Saavedra & Hubbard, 1991). This relation appears robust, remaining significant even when tested with a number of other factors such as teachers' receptivity to change, involvement in decision making, and age (Ringwalt et al., 2003). However, mixed findings exist. In a nationally representative sample of school-based prevention programs, Payne and Eckert (2010) showed that training quality was negatively associated with frequency and duration of program implementation. Technical assistance and coaching have also been linked to FOI. Similar to training, the presence of ongoing assistance as well as its perceived quality has been linked to higher levels of FOI (Mihalic & Irwin, 2003; Payne & Eckert, 2010; Ransford et al., 2009; Ringwalt et al., 2009). Coaching may be especially helpful in supporting implementation when it includes positive and corrective feedback, content-related information, and implementation tips (Rohrbach, Gunning, Sun & Sussman, 2010).

Establishing the link between professional development and implementation is essential to understand the extent to which teachers adopted the program. In relation to the present study, substantiating the link between training and teachers' implementation of program practices represents one step in evaluating the potential indirect effect of the *RC* approach on improvements in teacher–student interactions.

1.4.2. Fidelity of implementation and program outcomes

Implementers vary greatly in the degree to which they employ the practices of an adopted intervention (Durlak & Dupre, 2008). As such, FOI has received increased attention as researchers seek to measure and account for the degree of program use among study participants. Indeed, the positive relation between FOI and measured outcomes is well established in the intervention literature across various disciplines including education, community health, and psychology (Dane & Schneider, 1998; Durlak, 2010). Programs implemented with high levels of fidelity garner effect sizes two to three times greater, on average, compared to programs implemented with low levels of fidelity (Durlak & Dupre, 2008). Few studies, however, have assessed the link between FOI in SEL intervention and classroom outcomes such as teacher–student interactions. Exceptions include an evaluation of the Fast Track PATHS curriculum, which showed positive associations between FOI and classroom atmosphere, including cooperation, engagement, and teacher responsivity to students (CPPRG, 1999), and an evaluation of the Child Development Project, which demonstrated FOI related to improved teacher–student relationships and trust of teachers (Battistich, Schaps & Wilson, 2004). Our study contributes to work on SEL evaluations by including measures of FOI to determine (a) the extent to which teachers' use of the prescribed practices changed postintervention and (b) whether these changes led to improvements in classroom interactions.

1.4.3. Measuring fidelity of implementation

Implementation science is a developing field in which great strides in the assessment and measurement of FOI are being made. The present study contributes these efforts by addressing two limitations appearing in earlier work. First, many studies have relied on a single informant to report FOI, and this informant is often the one implementing the intervention (Domitrovich & Greenberg, 2000). As informants, implementers can reflect over large periods of time and report on the use of practices that are infrequent or difficult to observe. Further, asking implementers about their practices requires minimal resource expenditure. However, implementer responses may reflect social desirability or a biased perception of program use rather than an objective account of their implementation (Adams, Soumerai, Lomas & Ross-Degnan, 1999). Observational measures offer an alternative to teacher-reported measures and, in particular, may provide less biased reports of teachers' practices. However, observations also have limitations. For instance, conducting observations is expensive, observers can typically only be present in the classroom on a few days of the year, and observers may be unable to observe teachers' use of low frequency practices. In accordance with recommendations (Mowbray, Holter, Teague & Bybee, 2003), we employ both teacher-reported and observational measures, as they provide data to demonstrate convergent validity and help to compensate for bias that may be inherent when self-report or observational data collection is used in isolation.

A second limitation in past implementation measurement has been the exclusion of the control group (Century et al., 2010). Traditionally, researchers have typically assumed that “business as usual” in the control group equates to an absence of the intervention. However, given the number of programs present in schools, many of which share similar objectives, this assumption may be untenable (Greenberg, 2010). That is, one cannot assume that a business as usual provides an uncontaminated comparison

condition. Therefore, it is imperative to measure the presence of intervention practices irrespective of treatment status in order to quantify and account for the actual (rather than the idealized) difference in the presence of intervention components among groups (Hulleman & Cordray, 2009).

The *RC* approach offers a good illustration of the need to assess FOI in all participant groups. *RC* practices may be present in classrooms in the absence of teachers' formal exposure to the program for multiple reasons. First, *RC* practices are consistent with classic educational and developmental theory (e.g., that of Vygotsky's Social Development Theory (Vygotsky, 1978) and Bronfenbrenner's Ecological Systems Theory (Bronfenbrenner, 1979)) as well as contemporary educational research pointing to the importance of classroom environment and interactions on student learning and development (Pianta & Hamre, 2009). In essence, the 10 *RC* practices represent a bundle of teaching best practices. Therefore, it is plausible that teachers have been exposed to *RC*-like practices during their preservice training, through colleagues, or in other professional development. As an example, the *RC* approach emphasizes the importance of establishing classroom expectations and routines in the first weeks of school. Wong and Wong (2009) provide similar recommendations and guidelines in their popular book, *The First Days of School: How to Be an Effective Teacher*. In this way, teachers may be implementing components of the *RC* approach without even thinking of them as *RC* practices. Second, *RC* materials are widely disseminated. NEFC distributes a monthly *RC* newsletter and has published multiple books on the *RC* approach and its component practices. These products are accessible to teachers regardless of their training history with *RC*. Exposure to these materials may influence non-*RC* trained teachers to adopt aspects of the *RC* approach. Either of these scenarios could result in program contamination, pointing to the importance of measuring the presence of *RC* practices in all participating teachers' classrooms.

1.5. Other factors contributing to teacher–student interaction quality

American schools are highly variable in the practices offered to children, particularly in relation to teacher–student interaction quality. There are mixed findings regarding the extent to which specific structural features, such as class size, teachers' level of education, and teachers' years of teaching experience, relate to the observed quality of the emotional and instructional climates in elementary classrooms (NICHD ECCRN, 2002, 2005; Pianta, Belsky, Houts & Morrison, 2007; Stuhlman & Pianta, 2009). Some studies have found significant but weak associations, whereas other studies have found no significant relation. Given the inconsistency of these findings, we included these variables as covariates in our models in an attempt to increase the precision of coefficient estimates.

1.6. The present study

The experimental design provides an opportunity to test the impact of training in the *RC* approach on outcomes of interest. In the present study, these were teachers' implementation of *RC* practices and teacher–student interaction quality. In essence, tests of the direct effects of *RC* training provide "causal description" (Shadish, Cook & Campbell, 2002, p. 9). We hypothesized that teachers in the *RC* experimental schools (and thus receiving training) would demonstrate greater use of *RC* practices and higher quality interactions with students compared to the control group, after controlling for teacher and classroom characteristics and interaction quality assessed prior to intervention (Fig. 1, panel A). Given the possibility that intervention practices may be naturally prevalent in the control condition, or alternatively that teachers in the experimental group may not adopt *RC* practices, it is possible that *RC* training would show little impact on teachers' use of *RC* practices. Thus, understanding the extent to which the training led to greater use of *RC* practices is an important aspect of program evaluation in addition to the examination of the effects on teacher–student interactions.

In addition to making a causal description, we were also interested in the "causal explanation" (Shadish et al., 2002, p. 9) of the relation between training in the *RC* approach and interaction quality. We hypothesized that use of *RC* practices (operationalized as

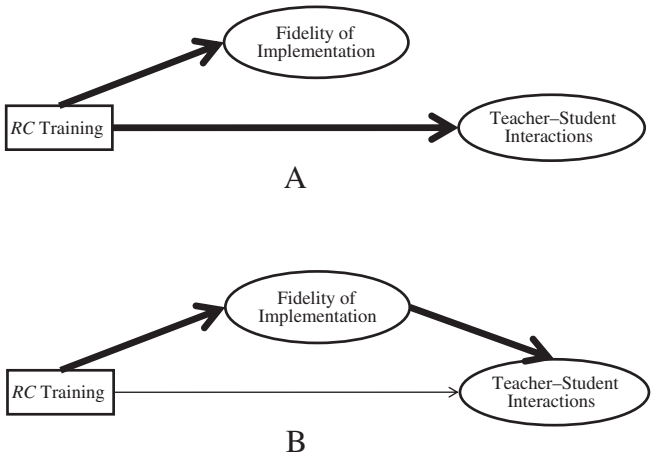

Fig. 1. Theoretical models of the direct (panel A) and indirect (panel B) effects of training in the *RC* approach on teacher–student interaction quality. Hypothesized paths are shown in bold. Model controls and observed indicators of latent variables are removed for clarity.

FOI) was the mechanism by which training in the *RC* approach would relate to improved teacher–student interaction quality. We expected that teachers trained in the *RC* approach would have higher FOI of *RC* practices compared to their control counterparts who did not receive training, which would in turn, relate to greater improvements in teacher–student interaction quality (Fig. 1, panel B).

2. Method

The present study was part of a larger three-year longitudinal cluster randomized control trial examining the impact of the *RC* approach on teacher–student interaction quality and student academic achievement as students progressed through third-, fourth-, and fifth-grade classrooms (Rimm-Kaufman et al., submitted for publication). We examine teacher–student interaction quality as the dependent variable of primary interest utilizing the first 2 years of cross-sectional data collected from third- and fourth-grade teachers.

2.1. School sample

Elementary schools from a single school district in the mid-Atlantic were recruited to participate in a randomized controlled trial of the *RC* approach. The district was in the process of mandating the adoption of behavior management programs in all schools, and the *RC* approach was offered as one option. Twenty-four schools were chosen for inclusion in the study based on their expressed interest in adopting the *RC* approach. School-level randomization was stratified based on proportion of student eligibility for free or reduced lunch and student racial/ethnic composition. Stratified randomization yielded 13 schools assigned to an *RC* experimental group and 11 schools assigned to the control group with the promise of receiving the intervention upon study completion (control schools did not adopt an interim alternative behavior management program). Assessment at randomization demonstrated that participating schools were demographically diverse in their representation of free or reduced lunch status (range = 2% to 72%, $M = 26\%$) and minority student composition (range 17% to 86%, $M = 55\%$). Independent sample *t*-tests revealed no significant differences between experimental and control schools on these variables, $ps > .05$. School size ranged from 289 to 986 students, and the percent of English Language Learners ranged from 5% to 75%, also with no significant differences between experimental and control schools on these characteristics, $p > .05$.

2.2. Teacher sample

The teacher sample included 239 third- and fourth-grade teachers (132 experimental, 107 control) from the participating schools. The number of teachers participating from each school ranged from 5 to 18 in the experimental schools and from 5 to 15 in the control schools. Eighty-eight percent of participating teachers were women. Approximately 85% of participating teachers reported being European American, 5% reported being African American, 3% reported being Hispanic/Latino, and 2% reported being Asian. Participating teachers were 39 years old, on average, with 10 years of teaching experience, and 65% reported holding a Master's degree. Specific values for experimental and control groups are presented in Table 1. Examination of these demographic characteristics showed no significant differences between the experimental and control teachers with the exception of race in which the intervention group had a higher proportion of teachers who were from racial/ethnic minority groups, $p < .05$.

2.3. Measures

2.3.1. Teacher–student interaction quality

The quality of teacher–student interactions was assessed using the Classroom Assessment Scoring System (CLASS; Pianta, LaParo & Hamre, 2008b). The CLASS is a measure of teacher–student interaction quality assessed on 10 dimensions categorized into three domains: Emotional Support, Classroom Organization, and Instructional Support. Each of the 10 dimensions is defined by specific behavioral indicators and scored on a 1 to 7 scale with a higher score indicating more positive interactions (i.e., 1 and 2 = *low*; 3, 4, and 5 = *mid*; and 6 and 7 = *high*). Emotional Support comprises four dimensions including Positive Climate, Negative Climate (reversed scaled), Teacher Sensitivity, and Regard for Student Perspectives. Emotional Support assesses the warmth and connection

Table 1
Demographic characteristics of teacher participants by treatment assignment.

Teacher characteristic	Experimental group ($n = 132$)				Control group ($n = 107$)			
	n	%	*M*	*SD*	*n*	%	*M*	*SD*
Women	114	86			97	91		
Average age			38.76	11.40			39.13	12.62
Ethnicity								
European American	102	80			94	90		
African American	11	9			2	2		
Hispanic/Latino	6	5			1	1		
Asian	4	3			2	2		
Other	5	4			2	2		
Years of teaching experience			9.26	6.74			10.98	9.41
Teacher has a Master's degree	59	63			56	67		

among teacher and students and between students, responsiveness to students' academic and social needs, and the degree to which autonomy and responsibility are encouraged in the classroom. Classroom Organization comprises three dimensions, Behavior Management, Productivity, and Instructional Learning Formats, and assesses the teacher's maximization of instructional time, ability to provide clear and consistent behavior management strategies that prevent and efficiently redirect misbehavior, and use of strategies to promote student interest and maximize opportunities for learning. Instructional Support comprises three dimensions, Concept Development, Quality Feedback, and Language Modeling. Instructional Support measures the teacher's promotion of higher-order thinking skills as well as the emphasis on understanding over rote knowledge, the extent to which the teacher provides specific and informative feedback, and the use of language modeling and facilitation techniques.

The theoretical three-factor structure has been supported by empirical psychometric work (Pianta & LaParo et al., 2008b). In other studies, researchers have compiled CLASS domain scores to create a single composite, given moderate to strong correlations among the CLASS domains (see Brown et al., 2010; Ponitz, Rimm-Kaufman, Grimm & Curby, 2009). This compilation of domain scores is similar to the approach taken in the present study in which domain scores serve as indicators for a latent Teacher–Student Interaction Quality variable. Bivariate correlations between the three domains as well as the factor loadings of the three domains on the Teacher–Student Interaction Quality factor, presented in the Results section, support our treatment of the CLASS as a unidimensional assessment.

Previous research using the CLASS has demonstrated evidence of the concurrent and predictive validity of CLASS scores. CLASS domains have been shown to correlate positively and significantly with widely used measure of early childhood classroom quality, the Early Childhood Environment Rating Scale — Revised (Harms, Clifford & Cryer, 2005), with correlation coefficients ranging from .45 to .63 (Pianta & LaParo et al., 2008b). Further, higher CLASS scores have been associated with growth in reading and mathematics skills during early elementary schooling as well as higher teacher ratings of social, language, and mathematics skills in kindergarten (Pianta, Belsky, Vandergrift, Houts, & Morrison, 2008a; Pianta et al., 2002).

At the conclusion of a 2-day training, 13 coders established initial inter-rater reliability on a minimum of 10, 15- to 20-minute master-coded classroom observations. In accordance with developers' scoring recommendations (Pianta & LaParo et al., 2008), inter-rater reliability was assessed using the percent agreement with the master code, plus or minus one scale point, which exceeded 80% for all coders. We continued to assess inter-rater reliability on an ongoing basis via semimonthly calibration meetings in which all coders watched and scored a 15-minute observation selected at random. Coders' scores were recorded and used to calculate intraclass correlations at six time points during the data coding period; values ranged from .73 to .85. Additionally, throughout the coding period, we conducted a randomized audit of raters' codes. Four to eight segments were selected randomly from each rater and were double-coded by a certified CLASS trainer. Averaging across the selected tapes, percent agreement within one point between the original coder and the master coder exceeded 80%.

Using video recordings, two 15-minute segments were scored from each of the five 60-minute classroom observations, resulting in 10 observations scored across 5 days. To minimize rater bias, each 15-minute segment was assigned randomly to an independent rater, blind to treatment condition, such that a coder never scored two segments from a given observation. Additionally, raters scored no more than 30% of each teacher's segments. CLASS domain scores were created by averaging ratings of the respective dimensions for each 15-minute segment. Subsequently, domain scores were averaged across the 10 segments yielding a single indicator of observed teacher–student interaction quality in each domain of Emotional Support, Classroom Organization, and Instructional Support. Measures of internal consistency for the individual domains, derived from our data, provide evidence of adequate reliability (Emotional Support $\alpha = .78$; Classroom Organization $\alpha = .67$; Instructional Support $\alpha = .86$).

2.3.2. Fidelity of implementation (observed)

Teachers' observed fidelity of implementation was assessed using the Classroom Practices Observation Measure (CPOM). The CPOM (Abry, Brewer, Nathanson, Sawyer & Rimm-Kaufman, 2010) is a 16-item measure used to assess teachers' observed adherence to *RC* practices in the classroom. The measure described practices without using specific *RC* terminology. Example items include, "Three to five general, positively worded rules are posted in the classroom," "Teacher asks questions or makes statements that invite students to remember expected behaviors," "Students make individualized choices related to an academic lesson or goal," and "Teacher facilitates students sharing brief, personal news or stories with the rest of the class." The 16-item version was administered during morning observations and an abbreviated 10-item version, which excluded items pertaining to the Morning Meeting, was used during mathematics observations. All items were coded on a 3-point Likert scale (*not at all characteristic*, *moderately characteristic*, and *very characteristic*) with higher scores indicating that the item was very characteristic during the observation. Two items were reversely scored.

Upon completion of a 2-day training, four coders established initial inter-rater reliability on a minimum of eight 60-minute master-coded videos using exact agreement, which exceeded 80% for all coders. Inter-rater reliability was evaluated on an ongoing basis through monthly calibration meetings in which coders viewed and coded a 60-minute video observation. These scores were recorded and used to calculate intraclass correlations at five time points during the coding period, which ranged from .74 to .88.

CPOM scores were based on the full 60-minute observation and were coded live, on site by one of four research assistants. For each observation, the 16 items were averaged to create a single score. These scores were then averaged across the five observations to create a single numeric indicator of teachers' observed fidelity of implementation of *RC* practices ($\alpha = .88$ in the current sample).

2.3.3. Fidelity of implementation (teacher-reported)

Teachers reported their perceived adherence to *RC* practices using the Classroom Practices Teacher Survey (CPTS). The CPTS (Nathanson et al., 2007a) is a 46-item measure assessing teachers' reported adherence to *RC* practices (e.g., "In the morning we have a

class meeting where we sit in a circle facing one another," and "When introducing new materials, I ask students to demonstrate ideas for how to use and care for the materials."). The teachers completed the survey at a single time point in the final 8 weeks of the school year via an online survey distributor. Teachers were asked to reflect over the course of the school year and to respond to each item on a 5-point Likert scale ranging from *not at all characteristic* to *extremely characteristic*. Seven items were reversely scored. To minimize bias of teachers' responses, items did not contain *RC* terminology. A single mean score was computed by averaging across all 46 items ($\alpha = .91$ in the current sample). This score served as an indicator of teachers' perceived implementation of *RC* practices.

A second teacher-reported measure of FOI, the Classroom Practices Frequency Survey (CPFS; Nathanson et al., 2007b) was administered concurrently with the CPTS. The CPFS is an 11-item measure assessing the frequency of teachers' use of *RC* practices on an 8-point Likert scale ranging from *almost never* to *more than once per day*. Example items include, "When a rule or procedure is introduced, I ask students to model what following the rule or procedure looks like," and "I provide opportunities for students to choose how to do work, what kind of work to do, or both." A single mean score was computed by averaging across all 11 items ($\alpha = .88$ in the current sample). This score served as the indicator of teachers' reported frequency of implementation of *RC* practices. This measure complimented the adherence measures by assessing FOI in a conceptually distinct domain. For example, a teacher that implements all the components of the Morning Meeting but holds the meeting only once a week would score high in adherence but low in frequency.

All three fidelity of implementation measures, the CPOM, CPTS, and CPFS, were created and refined specifically for use in the larger Responsive Classroom Efficacy Study (Rimm-Kaufman et al., submitted for publication). The second author and members of the research team consulted with NEFC staff throughout the measurement development process to help ensure that items represented the breadth of *RC* practices and were written in a way that would maximize the potential to capture subtle nuances between teachers adhering to *RC* principles and practices and teachers that may have adopted *RC*-like practices in their classrooms in the absence of formal training in the approach. In addition, the CPTS was piloted in read-aloud cognitive interviews (see Collins, 2003) with four elementary teachers prior to administration. Results from independent sample *t*-tests indicated that CPTS scores for control and experimental teachers did not differ at pretest, but subsequent to training, teachers in the intervention group had significantly higher CPTS scores on average, $p < .05$. Other research has supported the predictive validity of the CPTS, linking higher CPTS scores to higher mathematics achievement (Abry, Rimm-Kaufman, Hulleman, Thomas & Ko, 2012). In the present study, bivariate correlations between the three FOI measures ranged from .63 to .75 demonstrating a substantial proportion of shared variance across the measures (40% to 56%).

2.3.4. Classroom and teacher characteristics

Information on classroom and teacher characteristics, including whether the teacher had a Master's degree, years of teaching experience, and class size, was collected via online demographic questionnaires administered in the spring concurrently with the self-report assessments of FOI.

2.4. Procedures

In the fall of 2007, third- and fourth-grade teachers received a letter inviting participation and detailing the study components. The response rate was over 95% and consenting teachers received $100.00 for participating in observational data collection and completing surveys.

NEFC consulting teachers certified to conduct trainings in the *RC* approach led all trainings, and a local coordinator was put in place to help ensure the integrity of *RC* trainings conducted with the sample. However, fidelity data pertaining to the administration of the trainings were not collected. The *RC* 1 training, a weeklong training conducted in the summer, introduced specific *RC* practices pertinent to building classroom community and supporting classroom management. The second *RC* training, *RC* 2, lasting one week, expanded on the practices taught in the first training and introduced additional practices such as techniques for instruction, problem solving, and working with families. Approximately three in-person consultations with *RC* coaches followed each of the two *RC* trainings. In most situations, a teacher's initial trainer served as his or her coach during the school year. Coaching sessions were designed to support teachers in their use of the material presented in the most recent *RC* training attended. In these visits, coaches observed teachers and offered specific feedback and support for their implementation of *RC* practices. Face-to-face visits were supplemented with email and phone communication throughout the school year at the teachers' discretion. Coaches also disseminated short articles with timely ideas for teachers in the months that they were not onsite. In addition, each school was provided with a library of NEFC books and DVDs to support their adoption of the *RC* approach. Additionally, a local consulting teacher placed within the district facilitated trainings and coaches' visits.

Table 2 describes the schedule for data collection. Pretest measures of teacher–student interaction quality were collected in the spring of 2008. Third- and fourth-grade teachers in the experimental group received *RC* 1 training in summer 2008 and *RC* 2 training in summer 2009. Four research assistants conducted classroom observations (as described in the following paragraph) to assess teacher–student interaction quality and fidelity of implementation. Third-grade intervention and control teachers were observed during the 2008–2009 school year in their first year of implementation after the *RC* 1 training. Fourth-grade teachers were observed during the 2009–2010 school year in their second year of implementation after the *RC* 2 training. All third-grade teachers completed demographic questionnaires and self-assessments of their use of *RC* practices in the spring of 2009. Fourth-grade teachers completed the same assessments in the spring of 2010.

Project staff videotaped teachers on five "typical" school days spaced throughout the school year. Two observations were conducted during the first hour of the morning and three observations were conducted for one hour during mathematics instruction.

Table 2
Timeline of data collection and responsive classroom training for third- and fourth-grade teachers.

Teacher cohort	Spring 2008	Summer 2008	2008–09	Spring 2009	Summer 2009	2009–10	Spring 2010
Third-grade	Baseline	*RC* 1 training (treatment teachers only)	Classroom observations	Teacher surveys	*RC* 2 training (treatment teachers only)		
Fourth-grade	Baseline	*RC* 1 training (treatment teachers only)			*RC* 2 training (treatment teachers only)	Classroom observations	Teacher surveys

Third- and fourth-grade teachers in the control group received no exposure to *RC* training or coaching support and continued "business as usual" classroom instruction.

2.5. Analytic approach and data screening

We employed a structural equation modeling (SEM) approach to test the effects of *RC* training on use of *RC* practices and improvements in teacher–student interactions. Structural equation models allow researchers to test multiple hypothesized associations among observed and latent variables simultaneously. In SEM, latent variables (i.e., factors) are formed from observed variables referred to as indicators. The error terms associated with each indicator partition out both the unique random and systematic variance unexplained by the factor, yielding a latent variable with no associated measurement error (Kline, 2011), one of the benefits of SEM.

Scores from the three FOI measures, the CPTS, CPOM, and CPFS, were modeled as indicators of the latent Fidelity of Implementation variable. The three CLASS domain scores, Emotional Support, Classroom Organization, and Instructional Support were modeled as indicators of the latent Teacher–Student Interaction variable. As shown in Table 3, correlations between the FOI indicators were moderate ($r = .63$ to .75), as were correlations between the teacher–student interaction indicators ($r = .58$ to .70). In both cases, the moderately sized correlations suggest the three observed variables hypothesized to represent each underlying construct shared a significant proportion of their variance. Factor loadings assessed in the measurement model (presented in the Results section) further established the existence of a single underlying factor for each construct.

Table 3
Correlations and descriptive statistics for predictors, outcomes, and covariates.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Predictors and outcomes														
1. *RC* training (0 = no training)	–													
Fidelity of Implementation														
2. Classroom Practices Teacher Survey	.87*	–												
3. Classroom Practices Observation Measure	.90*	.66*	–											
4. Classroom Practices Frequency Survey	.85*	.75*	.63*	–										
Posttest Teacher–Student Interaction Quality														
5. Posttest Emotional Support	.20	.26*	.31*	.32*	–									
6. Posttest Classroom Organization	.12	.22*	.21*	.19*	.64*	–								
7. Posttest Instructional Support	−.16	.09	.14	.16*	.70*	.58*	–							
Covariates														
Pretest Teacher–Student Interaction Quality														
8. Pretest Emotional Support	−.01	.17	.10	.11	.55*	.35*	.41*	–						
9. Pretest Classroom Organization	−.04	.09	−.06	.04	.22*	.23*	.21*	.54*	–					
10. Pretest Instructional Support	−.09	.05	.01	−.01	.39*	.32*	.44*	.55*	.44*	–				
Other classroom and teacher variables														
11. Grade level (0 = third grade)	.04	−.02	.08	−.09	−.16*	−.08	−.36*	−.06	−.10	.06	–			
12. Master's degree (0 = no Master's degree)	−.04	−.04	.03	−.04	−.02	.00	−.15	.00	−.05	−.01	.04	–		
13. Years of teaching experience	−.27	−.01	−.19	−.07	−.13	.04	−.08	−.03	.09	.19	.02	−.03	–	
14. Class size	−.15	−.12	−.08	−.12	−.12	−.07	−.22*	.05	.06	−.20	−.03	.05	.09	–
N	24[a]	181	179	180	181	181	181	132	132	132	233	178	178	178
M	.52	3.67	1.61	4.85	5.10	5.81	2.79	5.23	5.90	3.12	.54	.65	10.07	22.65
SD	.51	0.58	0.31	1.50	0.41	0.40	0.49	0.46	0.44	0.60	.50	.48	8.13	5.07
Min	0	2.26	1.02	1.45	3.69	4.23	1.42	4.00	4.17	1.83	0	0	1	0
Max	1	4.74	2.30	7.91	6.18	6.50	4.33	6.38	6.67	4.75	1	1	38	40

Note.
[a] = School-level variable.
* $p < .05$.

Standard assumptions of linearity, homoscedasticity, and normality were assessed through visual examination of histograms, residual plots, and quantile–quantile (Q–Q) plots of the residuals produced by the component regression models. Consistent with the assumption of linearity, standardized residual values appeared uncorrelated with standardized predicted values of the outcomes as evidenced by a lack of curvature in the plots. Supporting the assumption of homoscedasticity, standardized residual values were equally distributed across predicted values of the outcomes. Q–Q plots demonstrated normal distribution of the residuals. Finally, histograms revealed that individual variables were also normally distributed.

We estimated all models in Mplus (Muthén & Muthén, 1998–2009) using full information maximum likelihood (FIML) estimation. All models were specified using TYPE = COMPLEX to account for the nonindependence of observations (i.e., teachers nested within schools) when computing standard errors and chi-square values. Intraclass correlations for model outcomes (72% for the latent FOI variable and 18% for the latent Teacher–Student Interaction Quality variable) indicated a substantial amount of school-level variance, supporting the need to adjust for clustering effects. TYPE = COMPLEX adjusts for nonindependence among cases while maintaining a single level of analysis, in this case, the teacher-level. We did not employ a two-level model because the small number of schools in combination with the large number of parameters being estimated led to difficulties with model convergence when using FIML estimation. Two-level models using a Bayesian estimator confirmed the TYPE = COMPLEX results but are not presented here because they do not yield readily interpretable indices of model fit. Model fit was assessed using four indices: chi-square difference testing, the comparative fit index (CFI), the Tucker Lewis Index (TLI), and the root mean square error of approximation (RMSEA). CFI and TLI values greater than .90 and RMSEA values less than .08 indicate satisfactory model fit (Bentler, 1990; Bentler & Bonett, 1980; Browne & Cudeck, 1992).

Of the 239 teachers, 58 (24%) had missing CPTS scores, 60 (25%) had missing CPOM scores, 59 (25%) had missing CPFS scores, and 58 (24%) had missing posttest teacher–student interaction quality data. These teachers left the sample during the time between the collection of the pretest and posttest measures of interaction quality (a 1- and 2-year span for third- and fourth-grade teachers, respectively). Comparisons of pretest measures of interaction quality between intervention and control teachers for those remaining in the sample did not significantly differ ($p > .05$), suggesting that attrition did not compromise the integrity of the randomization. The probability that an outcome was missing was associated with teachers' grade level and teachers' observed implementation. These variables were included in the analyses, and we could find no other plausible missing data mechanism; thus, the data were presumed to be missing at random. To minimize bias in parameter estimates, FIML was used in the estimation step of the analysis (Enders, 2001), which has the added benefit of preserving the original sample size (Little & Rubin, 1987).

3. Results

First, we present the results of the measurement model. Second, we present results of the two SEM structural models that tested our research questions. Standardized factor loadings and parameter estimates and model fit indices are presented in Table 4. An alpha of .05 was used to evaluate all tests of statistical significance.

3.1. Measurement model

The initial measurement model included three latent variables (FOI, pretest Teacher–Student Interaction Quality, and posttest Teacher–Student Interaction Quality) and five exogenous observed variables: a dichotomous treatment condition variable (0 = control [no *RC* training], 1 = experimental [*RC* training]), a dichotomous grade level variable (0 = third grade, 1 = fourth grade), a dichotomous indicator of whether teachers had a Master's degree (0 = no Master's degree, 1 = Master's degree), teachers' years of teaching experience, and teachers' class size. Based on a test of measurement invariance, factor loadings for the pretest and posttest measures of Teacher–Student Interaction Quality were constrained to be equal. Additionally, residual variances for like indicators of the pretest and posttest Teacher–Student Interaction Quality latent variables were allowed to correlate (Marsh & Hau, 1996). Pretest Teacher–Student Interaction Quality, grade level, receipt of a Master's degree, years of teaching experience, and class size were hypothesized as model controls. However, contrary to our expectation, receipt of a Master's degree, teaching experience, and class size did not significantly relate to the FOI or Teacher–Student Interaction Quality factors, and thus they were dropped to ensure parsimony of the model. The final measurement model (Table 4, Model 1) showed adequate fit (CFI = .96, TLI = .93, RMSEA = .06 [90% CI = .04 to .09]). Standardized factor loadings, all greater than .73 and statistically significant ($p < .05$), revealed satisfactory contribution of each indicator to its underlying factor. Latent variables correlated with each other and treatment assignment (i.e., *RC* training) in the expected directions: teachers trained in the *RC* approach tended to show greater levels of FOI; FOI was positively related to posttest Teacher–Student Interaction Quality; and pretest Teacher–Student Interaction Quality was positively related to FOI and posttest Teacher–Student Interaction Quality suggesting teachers high in pretest Interaction Quality showed greater use of *RC* practices and had higher subsequent Interaction Quality. Unexpectedly, training in *RC* did not significantly relate to Teacher–Student Interaction Quality at posttest.

3.2. Structural models

After specifying the measurement model, we tested the direct effects of training in the *RC* approach on FOI and posttest Teacher–Student Interaction Quality, controlling for teachers' grade level and pretest Interaction Quality (Table 4, Model 2). This structural model contained six direct effects: *RC* training predicting FOI and posttest Teacher–Student Interaction Quality; pretest Interaction Quality predicting FOI and posttest Teacher–Student Interaction Quality; and teachers' grade level predicting FOI and

Table 4
Standardized parameter estimates for factor loadings and path coefficients, and fit indices for measurement and structural models.

	Model 1	Model 2	Model 3 (final)
Factor loadings			
Posttest Teacher–Student Interaction Quality → Emotional Support	.88*	.89*	.89*
Posttest Teacher–Student Interaction Quality → Classroom Organization	.74*	.76*	.76*
Posttest Teacher–Student Interaction Quality → Instructional Support	.81*	.82*	.81*
Fidelity of Implementation → Classroom Practices Teacher Survey	.85*	.86*	.85*
Fidelity of Implementation → Classroom Practices Observation Measure	.82*	.80*	.81*
Fidelity of Implementation → Classroom Practices Frequency Survey	.82*	.83*	.82*
Path coefficients			
RC training → Fidelity of Implementation	–	1.62*	1.62*
RC training → Posttest Teacher–Student Interaction Quality	–	0.18	−0.68*
Fidelity of Implementation → Posttest Teacher–Student Interaction Quality	–	–	0.52*
Pretest Teacher–Student Interaction Quality → Posttest Teacher–Student Interaction Quality	–	0.49*	0.31*
Pretest Teacher–Student Interaction Quality → Fidelity of Implementation	–	0.28*	0.21*
Grade level → Fidelity of Implementation	–	−0.06	−0.06
Grade level → Posttest Teacher–Student Interaction Quality	–	−0.23*	−0.20*
RC training → Fidelity of Implementation → Posttest Teacher–Student Interaction Quality	–	–	0.85*
Model fit statistics			
χ^2	72	87	77
df	36	40	39
CFI	.96	.94	.95
TLI	.93	.92	.93
RMSEA (90% CI)	.06 (.04–.09)	.07 (.05–.09)	.06 (.04–.08)

Note. CFI = comparative fit index; TLI = Tucker Lewis index; RMSEA (90% CI) = root mean square error of approximation (90% confidence interval).
* $p < .05$.

posttest Teacher–Student Interaction Quality. Results revealed that teachers trained in the *RC* approach had greater levels of FOI ($\beta = 1.62$, $p < .001$), though training in the *RC* approach did not significantly predict Teacher–Student Interaction Quality at posttest. Pretest Teacher–Student Interaction Quality significantly predicted FOI and posttest Interaction Quality. Fourth-grade teachers had significantly lower posttest Teacher–Student Interaction Quality compared to third-grade teachers.

Next, we tested a second structural model (Table 4, Model 3) in which we assessed the indirect effect of training in the *RC* approach on Teacher–Student Interaction Quality through teachers' use of *RC* practices by adding a path from FOI to posttest Teacher–Student Interaction Quality (see Fig. 2). This model necessitated the estimation of two additional parameters: FOI → posttest Teacher–Student Interaction Quality, and *RC* training → FOI → posttest Teacher–Student Interaction Quality. We retained pretest Interaction Quality and teachers' grade level as model controls, which maintained their effects on FOI and posttest Interaction Quality shown in Model 2. In Model 3, training in the *RC* approach predicted significantly greater FOI, $\beta = 1.62$, $p < .001$, which in turn predicted significantly greater improvements in posttest Teacher–Student Interaction Quality, $\beta = 0.52$, $p = .001$. Multiplying these two parameter estimates yielded the size of the indirect effect of training in the *RC* approach on Interaction Quality via FOI, $\beta = 0.85$, $p = .002$, which indicated Teacher–Student Interaction Quality increased by nearly three-quarters of a standard deviation through FOI, as a result of training in the *RC* approach. We calculated a bootstrapped confidence interval for the indirect effect parameter using 20,000 repetitions (95% CI = 0.31 to 1.94) to account for the nonnormality resulting from the multiplication of two standardized parameter estimates (Selig & Preacher, 2008). The confidence interval did not contain zero and verified the statistical significance of the indirect effect. In the presence of the significant positive indirect effect, a significant negative direct effect emerged between *RC* training and posttest Teacher–Student Interaction Quality, $\beta = -0.68$, p .03.

3.3. Final model

We compared the two structural models (Table 4, Model 2 and Model 3) to determine which model provided the best fit of the data. Fit indices indicated by CFI, TLI, and RMSEA values did not vary greatly between the two models (see the Model fit statistics section of Table 4). However, results of a chi-square difference test comparing these two models demonstrated that Model 3 had significantly better fit than Model 2, $\Delta\chi^2 = 10$, $\Delta df = 1$, $p = .002$. Thus, Model 3 was chosen as the more accurate representation of the data (Fig. 2). This model explained 69% of the variability in teachers' use of *RC* practices (i.e., FOI) and 32% of the variability in posttest Teacher–Student Interaction Quality.

4. Discussion

The objective of the present study was to examine the relation between randomized training in the *RC* approach, teachers' use of *RC* practices, and the quality of teacher–student interactions. Three findings emerged from our structural equation modeling. The first structural model (Model 2) showed a significant direct effect of *RC* training on teachers' use of *RC* practices, consistent with our hypothesis. Contrary to our hypothesis, there was no direct effect of *RC* training on improvements in the quality of teachers'

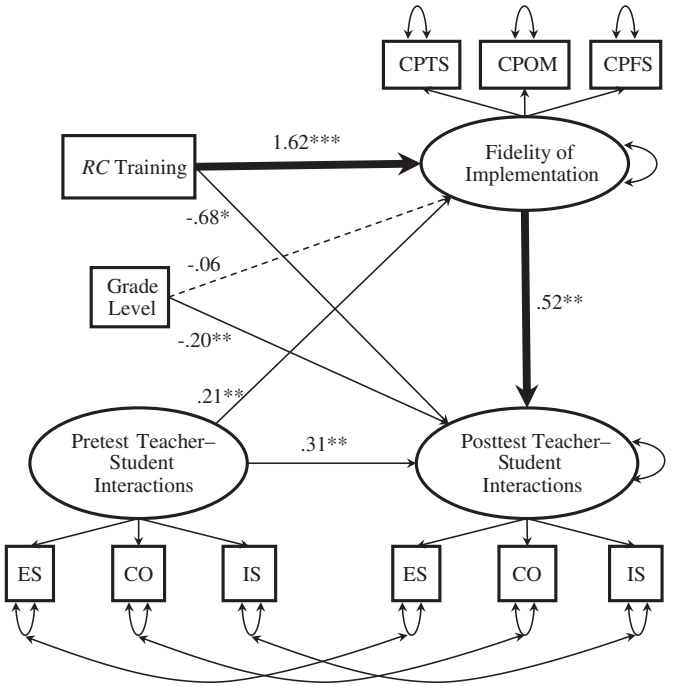

Fig. 2. Final model (Model 3) with training in the *RC* approach indirectly affecting teacher–student interaction quality through teachers' use of *RC* practices (FOI). Standardized parameter estimates are shown. Observed indicators are represented by squares and latent variables are represented by ovals. Solid paths indicate significance. Dashed paths indicate nonsignificance. Bolded paths indicate the indirect effect. Model factor loadings and fit statistics are shown in Table 3. CPTS = Classroom Practices Teacher Survey; CPOM = Classroom Practices Observation Measure; CPFS = Classroom Practices Frequency Survey; ES = Emotional Support; CO = Classroom Organization; IS = Instructional Support.

interactions with students. Model 3, the better fitting model, revealed a significant indirect effect of training in the *RC* approach on interaction quality via teachers' implementation of *RC* practices: teachers trained in the *RC* approach implemented more *RC* practices, which in turn was associated with greater improvements in interaction quality.

Despite the acknowledged importance of FOI, researchers have found treatment effects for SEL programs for classroom level outcomes (Brown et al., 2010; CPPRG, 1999). We did not replicate such impact findings. Thus, *RC* training alone did not produce improvements in teacher–student interactions, supporting the importance of examining FOI to better understand the null direct effect. Our finding that *RC* teachers had higher levels of fidelity than teachers in the control group indicates that the lack of a direct effect of *RC* training on teacher–student interaction quality was not a function of teachers' failure to uptake *RC* practices. The link between training in the *RC* approach and FOI is consistent with other work demonstrating the link between comprehensive, in-person training and technical support with fidelity to program components (Durlak & Dupre, 2008; Mihalic & Irwin, 2003; Ringwalt et al., 2003; Rohrbach, Grana, Sussman & Valente, 2006). In addition, the effect of *RC* training on FOI demonstrates that in the context of a school level initiative, time invested in professional development designed to instruct teachers in new approaches to discipline, organization, and instruction can produce changes in teacher practices, as measured by observation and self-report.

The indirect effect of *RC* training on improvements in teacher–student interaction quality replicates a small but growing body of work examining the link between the use of SEL interventions and classroom processes such as teacher–student interactions (Brown et al., 2010; CPPRG, 1999). SEL practices are designed to create well-structured classrooms where students are respected and provided an environment ripe for social and academic growth. Our analyses lend empirical evidence to support the conceptual link between SEL programming and healthy classroom environments. Further, the importance of FOI in our model is consistent with research on SEL and other school-based interventions demonstrating the importance of fidelity of implementation to ensuring program outcomes. In a recent meta-analysis of SEL interventions, Durlak et al. (2011) showed that SEL programs reporting implementation problems were less likely to have positive impacts on children's outcomes including social skills, social behaviors, and achievement.

The negative direct effect of training in the *RC* approach on interaction quality in the presence of an indirect effect via FOI was unexpected. This finding is consistent with what MacKinnon, Fairchild and Fritz (2007) describe as inconsistent mediation and suggests two opposing processes explaining teachers' improvements in interaction quality. The first process we have identified as FOI, yet the second process is unidentified.

It is possible that a teacher trained in the *RC* approach but that does not adopt *RC* practices is more likely to show little improvement in the quality of classroom interactions from year to year (i.e., a selection bias). A second explanation is the "implementation dip," defined as a decrease in performance as teachers try to change their practices (Fullan, 2001). When new programs require that teachers implement new practices, it frequently necessitates that they adapt or discard their familiar

approaches to teaching and interacting with students. Often, this process of change creates disequilibrium that may disrupt the typical flow of the classroom and cause teachers to feel uncertain or incompetent with their practice (Evans, 1996). In the present study, it is possible that some *RC* teachers experienced the training as a disturbance to their regular instruction. As one example, a teacher may have tried to implement the *RC* practice of Morning Meeting but instead facilitated a class meeting that blended their traditional approach with their new learning. This mix of practice could result in tension, disrupted flow, and fewer socially and emotionally supportive interactions among the teachers and students (assessed here as lower Emotional Support and Classroom Organization). Yet another explanation may be that an *RC* trained teacher may have high levels of fidelity, contributing the improvement of classroom interactions, while simultaneously enacting a behavior that has a negative influence on interaction quality.

Further work is necessary to understand the precise nature of the negative direct effect in the presence of FOI. Given the importance of FOI in the present study, we turn our attention toward factors not assessed in this paper that may contribute to teachers' variability in program uptake as well as the contextual factors that may support program implementation.

4.1. Explaining variability in implementation

Teachers trained in the *RC* approach demonstrated more use of *RC* practices compared to their control counterparts, yet FOI was still highly variable in the intervention group. Given comprehensive and interactive training, technical assistance in the form of ongoing coaching, the presence of a program liaison to assist with day-to-day issues, and access to an array of books and other *RC* materials, the question emerges, why did some teachers fail to uptake the practices introduced with the intervention? We offer three possible explanations.

One possible factor in implementation variability among the experimental group is training dosage (Dusenbury et al., 2003). Third-grade teachers in *RC* schools (approximately half of the teachers in the experimental group) received *RC* 1, a 1-week institute focusing on approximately half of the 10 *RC* practices. Fourth-grade teachers (the other half of the treatment sample) received *RC* 1 and *RC* 2, exposing them to *RC* practices not emphasized in the first training and affording them a second year of implementation and coaching prior to data collection.

A second possibility pertains to the existence of contextual factors that may lead to noncompliance. Positive perceptions of school climate, administrative support, and peer support have been shown to positively influence teachers' implementation of SEL interventions (Dariotis, Bumbarger, Duncan & Greenberg, 2008; Elias, Zins, Graczyk & Weissberg, 2003; Kam, Greenberg & Walls, 2003; Wanless, Patton, Rimm-Kaufman & Deutsch, 2012). At the same time, negative climate and lack of administrative support have been shown to have adverse effects on implementation (Kramer, Laumann & Brunson, 2000; Wanless et al., 2012). Anecdotally, teachers in our sample reported varied perceptions of principal support for the *RC* approach. For example, in some schools the administration financed the training of additional staff in the *RC* approach. In other schools, principals were willing to rework daily schedules to support the implementation of Morning Meeting on a daily basis. Other teachers reported that their use of *RC* practices suffered because of the administration's reluctance to integrate the approach school-wide and to provide them with helpful resources such as books and videos.

As a third explanation, personal factors including teachers' own attitudes and beliefs such as program buy-in are also likely to influence the extent to which intervention practices are adopted (Beets et al., 2008; Ringwalt et al., 2003; Rohrbach, Graham & Hansen, 1993). Teachers' buy-in may be greatest when there is no external mandate involved (Smylie, 1988), when teachers are involved in the decision making process surrounding program adoption (Parcel et al., 1991), and when the professional development appeals to teachers without being overwhelming (Kent, 2004). As described in the Method section, all schools in the district from which our sample was drawn were in the process of adopting behavior management programs, in essence, an external mandate. Perhaps as a result, nearly 100% of teachers in the intervention schools participated in the *RC* trainings. This compliance rate is unusually high and suggests that some teachers may have enrolled in training for reasons other than an inherent interest in learning to use the *RC* practices. Teachers may have felt compelled to attend training based on their principal's presentation of the circumstances. For these teachers, program buy-in may have been minimal, ultimately negatively affecting their implementation of *RC* practices.

4.2. Considerations for school personnel in the support of implementation

Variability in program uptake is inevitable. Few studies have reported 100% fidelity of implementation among all implementers, sites, or both (Durlak & Dupre, 2008). Moreover, our results suggest that SEL programs may have both positive (e.g., FOI) and negative mechanisms by which they influence classroom processes such as teacher–student interactions. Widespread variability in FOI combined with its potential to be a driver for positive program outcomes speaks to the importance of providing supports for implementers.

School psychologists, administrators, and other staff play an important role in selecting and supporting the adoption of interventions designed to improve the school and classroom climate. In fact, these key personnel are frequently the main source for what Fixsen, Blasé, Naoom and Wallace (2009) refer to as implementation core components. Implementation core components, or *implementation drivers* (Fixsen et al., 2009), include training in the intervention, ongoing coaching and consultation, implementer assessment, data systems used to guide decision making, administrative facilitation of program implementation, and systems coordination aimed to maintain the financial and human capital necessary to sustain the program. These implementation drivers interact to influence the success of program implementation and provide multiple mechanisms through which school psychologists and administrators can support implementation.

For example, armed with the knowledge that quality training and technical assistance are crucial to support intervention implementation and sustainability (Ransford et al., 2009; Ringwalt et al., 2009; Rohrbach et al., 2010), school officials tasked with

selecting programs must actively seek those providing adequate initial training as well as ongoing support. Failing to do so may result in fragmented and ineffective service. Schools should also be prepared to assess implementers and to use the data to inform decisions about implementation support and program continuation. Collecting data on implementation as well as targeted outcomes will help alert school personnel to the presence of potential unintended negative side effects. Further, organizational readiness, commitment, and compatibility with the goals of an intervention should be a consideration in program selection because it will influence buy-in and implementation (Greenberg, 2010; Rogers, 2002). By involving teachers in needs assessments and the process of intervention selection, administrators can promote support for the program and minimize resistance. Once a program is selected, a well-respected, internal program coordinator can help maintain commitment among implementers by maintaining open lines of communication and troubleshooting issues as they arise (McLaughlin & Mitra, 2001).

4.3. Limitations

Several limitations require mention. First, despite utilizing data from a randomized control trial, we cannot make causal attributions about the indirect effect of *RC* on improvements in teacher–student interaction quality via fidelity of implementation. In order to do so, the predictor as well as the intervening variable would need to be randomly assigned (Fairchild & McQuillin, 2010). In the present study, only treatment assignment (i.e., *RC* training) was manipulated experimentally. Thus, the path from *RC* training to FOI represents a causal "impact" finding, whereas the path from fidelity to interaction quality is only predictive in nature. Our findings do, however, provide strong evidence that implementation of *RC* practices is a mechanism by which *RC* training contributes to improvements in the quality of teacher–student interactions.

Second, we analyzed data from a sample of teachers receiving different dosages of training in the *RC* approach. This design may not have been ideal from the perspective of looking at teachers with the highest levels of implementation. However, NEFC reports that well over half of teachers trained in the *RC* approach only receive *RC* 1. Thus, the heterogeneity of the training increases the external validity of our findings and provides potentially useful information to other districts adopting the *RC* approach.

Third, although strategies were implemented to ensure the integrity and quality of the *RC* trainings and coaching consultations, we do not have quantitative data to ensure that these measures were effective. Nor do we have data to assess teachers' engagement or satisfaction with the coaching sessions. We acknowledge this limits the extent to which we can offer definitive explanations of the variability in FOI among *RC* teachers.

4.4. Future directions

The present study raises several new questions for future inquiry that will aid in our understanding of the effects of *RC* approach and contribute to the research base on SEL programming. For example, with the growing body of research describing the importance of high quality teacher–student interactions for improving children's social and academic competence (e.g., Pianta & Hamre, 2009), further work examining the effect of the *RC* approach on children's social and academic learning is warranted. Second, there is a need to examine the "active ingredients" of the *RC* approach in order to understand the extent to which certain practices have greater implications for targeted outcomes compared to others. Third, the evidential basis for the *RC* approach could be strengthened with knowledge about whether achieving a certain threshold of FOI is sufficient to produce change in classrooms and youth outcomes (McCartney, Bub & Burchinal, 2006).

A separate but related line of inquiry will assist with our understanding of how to support high levels of FOI. For instance, there is a need to situate research on the *RC* approach in the context of the process of teacher change. Teachers' implementation of the approach is a gradual process susceptible to vagaries associated with teachers' own personal development, their attitude toward the intervention, and the extent to which implementation of the intervention is supported by the school and community (Rimm-Kaufman & Hamre, 2010; Smylie, 1988; Wanless et al., 2012). Understanding the interrelation of these factors will lend insight into why implementation of the intervention practices is so uneven among teachers and shed light on ways to create conditions supportive of implementation. Additionally, it will be useful to examine FOI at different levels of the intervention, including the implementation drivers (Fixsen et al., 2009) such as the training and coaching sessions, as well as teachers' responsiveness to training and coaching. An expanded look at FOI may further explain variability in teachers' implementation of *RC* practices in the classroom and may elucidate ways to enhance teachers' introduction to the intervention as well as the ongoing supports designed to promote teachers' implementation of the program.

4.5. Conclusion

Federal policies aiming to improve access to high quality classrooms for all children prompt schools to adopt evidence-based SEL interventions designed to improve the classroom climate and children's behavioral and academic skills. Consequently, it is important to evaluate the extent to which such programs effectively enhance targeted outcomes. Equally important is to understand the processes underlying these changes. This study contributes to the small body of literature examining the impact of SEL programs on classroom processes, one of the mechanisms by which SEL interventions are presumed to improve youth outcomes.

Our findings have implications for both program evaluators and practitioners. First, our results echo the value of implementation data and demonstrate the importance of collecting and examining implementation data in both experimental and control groups. Without data from all participants, it is impossible to unpack a null overall effect such as the one between *RC* training and teacher–student interaction quality. Second, our findings suggest that SEL interventions, like the *RC* approach, hold potential to improve the

emotional, managerial, and instructional interactions teachers engage in with their students. However, in the context of an external mandate, providing professional development in an intervention may not be sufficient to induce changes in classroom processes. Contextual and personal factors will influence fidelity of implementation, which is essential for program effectiveness. Moreover, participating in professional development may even be detrimental to teacher–student interactions for teachers with lower levels of implementation. Taken together, these results are in accordance with extant work pointing to the importance of the combination of training and implementation to garner improvements in targeted outcomes (Haynes, Emmons & Woodruff, 1998), and they point to the importance of understanding the ways in which school officials can support teachers' implementation of adopted interventions.

References

Abry, T., Brewer, A., Nathanson, L., Sawyer, B., & Rimm-Kaufman, S. E. (2010). *Classroom Practices Observation Measure (CPOM).* Unpublished instrument, University of Virginia http://www.socialdevelopmentlab.org/wp-content/uploads/2012/05/CPOM-for-SDL-Website.pdf (Retrieved from).

Abry, T., Rimm-Kaufman, S. E., Hulleman, C. S., Thomas, J. B., & Ko, M. (2012, March). The how and for whom of program effectiveness: Dissecting the Responsive Classroom approach in relation to academic achievement. *Paper presented at the semi-annual meeting of the Society for Research on Educational Effectiveness, Washington D.C.*

Adams, A. S., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care*, *11*, 187–192, http://dx.doi.org/10.1093/intqhc/11.3.187.

Battistich, V., Schaps, E., & Wilson, N. (2004). Effects of an elementary school intervention on students' "connectedness" to school and social adjustment during middle school. *The Journal of Primary Prevention*, *24*, 243–262, http://dx.doi.org/10.1023/B:JOPP.0000018048.38517.cd.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K. K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science*, *9*, 264–275, http://dx.doi.org/10.1007/s11121-008-0100-2.

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246, http://dx.doi.org/10.1037/0033-2909.107.2.238.

Bentler, P., & Bonett, D. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606, http://dx.doi.org/10.1037/0033-2909.88.3.588.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. C. Whittrock (Ed.), *Handbook of research on teaching* (pp. 328–375) (3rd ed.). Washington, DC: American Education Research Association.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs program. *Journal of Educational Psychology*, *102*, 153–167, http://dx.doi.org/10.1037/a0018160.

Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, *21*, 230–258, http://dx.doi.org/10.1177/0049124192021002005.

Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation*, *31*, 1–17, http://dx.doi.org/10.1177/1098214010366173.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*, 229–238, http://dx.doi.org/10.1023/A:1023254226592.

Conduct Problems Prevention Research Group (1999). Initial impact of the fast track prevention trial for conduct problems II: Classroom effects. *Journal of Consulting and Clinical Psychology*, *67*, 648–657, http://dx.doi.org/10.1037/0022-006X.67.5.648.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimental design and analysis issues for field settings.* Chicago, IL: Rand McNally College Publishing.

Cook, T. D., Murphy, R. F., & Hunt, H. D. (2000). Comer's school development program in Chicago: A theory-based evaluation. *American Educational Research Journal*, *37*, 535–597, http://dx.doi.org/10.3102/00028312037002535.

Cordray, D. S., & Pion, G. M. (2006). Treatment strength and integrity: Models and methods. In R. R. Bootzin, & P. E. McKnight (Eds.), *Strengthening research methodology: Psychological measurement and evaluation* (pp. 103–124). Washington, DC: American Psychological Association.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23–45, http://dx.doi.org/10.1016/S0272-7358(97)00043-3.

Dariotis, J. K., Bumbarger, B. K., Duncan, L. G., & Greenberg, M. T. (2008). How do implementation efforts relate to program adherence? Examining the role of organizational, implementer, and program factors. *Journal of Community Psychology*, *36*, 744–760, http://dx.doi.org/10.1002/jcop.20255.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, *11*, 193–221 (Retrieved from http://tinyurl.com/3w7yrmf)

Durlak, J. A. (2010). The importance of doing well in whatever you do: A commentary on the special section, "implementation research in early childhood education". *Early Childhood Research Quarterly*, *25*, 348–357, http://dx.doi.org/10.1016/j.ecresq.2010.03.003.

Durlak, J. A., & Dupre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–350, http://dx.doi.org/10.1007/s10464-008-9165-0.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432, http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*, 237–256, http://dx.doi.org/10.1093/her/18.2.237.

Elias, M. J., Zins, J. E., Graczyk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling up of social–emotional and academic innovations in public schools. *School Psychology Review*, *32*, 303–320.

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, *36*, 103–112, http://dx.doi.org/10.1207/S15326985EP3602_5.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling*, *8*, 128–141, http://dx.doi.org/10.1207/S15328007SEM0801_7.

Evans, R. (1996). *The human side of school change: Reform, resistance, and the real-life problems of innovation.* San Francisco, CA: Jossey-Bass.

Fairchild, A. J., & McQuillin, S. D. (2010). Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of School Psychology*, *48*, 53–84, http://dx.doi.org/10.1016/j.jsp.2009.09.001.

Fixsen, D. L., Blasé, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, *19*, 531–540, http://dx.doi.org/10.1177/1049731509335549.

Fullan, M. G. (2001). *The new meaning of educational change.* New York, NY: Teachers College Press.

Greenberg, M. T. (2010). School-based prevention: Current status and future challenges. *Effective Education*, *2*, 27–52, http://dx.doi.org/10.1080/19415531003616862.

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale — Revised edition.* New York: Teachers College Press.

Haynes, N. M., Emmons, C. L., & Woodruff, D. W. (1998). School development program effects: Linking implementation to outcomes. *Journal of Education for Students Placed at Risk*, *3*, 71–85, http://dx.doi.org/10.1207/s15327671espr0301_6.

Howes, C. (2000). Social–emotional classroom climate in child care, child–teacher relationships and children's second grade peer relations. *Social Development*, *9*, 191–204, http://dx.doi.org/10.1111/1467-9507.00119.

Hughes, J. N., Luo, W., Kwok, O. M., & Loyd, L. K. (2008). Teacher–student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology*, *100*, 1–14, http://dx.doi.org/10.1037/0022-0663.100.1.1.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*, 88–110, http://dx.doi.org/10.1080/19345740802539325.
Kam, C. M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science*, *4*, 55–63, http://dx.doi.org/10.1023/A:1021786811186.
Karachi, T. W., Abbott, R. D., Catalano, R. F., Haggerty, K. P., & Fleming, C. B. (1999). Opening the black box: Using process evaluation measures to assess implementation and theory building. *American Journal of Community Psychology*, *27*, 711–731, http://dx.doi.org/10.1023/A:1022194005511.
Kent, A. M. (2004). Improving teacher quality through professional development. *Education*, *124*, 573–584 (Retrieved from http://tinyurl.com/4ull4x4)
Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*, 262–273, http://dx.doi.org/10.1111/j.1746-1561.2004.tb08283.x.
Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
Kramer, L., Laumann, G., & Brunson, L. (2000). Implementation and diffusion of the Rainbows program in rural communities: Implications for school-based prevention programming. *Journal of Educational and Psychological Consultation*, *11*, 37–64, http://dx.doi.org/10.1207/S1532768XJEPC1101_4.
Kress, J. S., & Elias, M. J. (2006). School-based social and emotional learning programs. In K. A. Renninger, & I. E. Sigel (Eds.), *Handbook of child psychology* (pp. 592–618) (6th ed.). Hoboken, NJ: Wiley, http://dx.doi.org/10.1002/9780470147658.chpsy0415.
Little, R., & Rubin, D. (1987). *Statistical analysis with missing data* (2nd ed.). New York, NY: Wiley.
MacKinnon, D. P., Fairchild, A. S., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614, http://dx.doi.org/10.1146/annurev.psych.58.110405.085542.
Marsh, H. W., & Hau, K. T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education*, *64*, 364–390 (Retrieved from http://www.jstor.org/stable/20152499)
McCartney, K., Bub, K., & Burchinal, M. (2006). Selection, detection, and reflection. *Monographs of the Society for Research in Child Development*, *71*, 105–126, http://dx.doi.org/10.1111/j.1540-5834.2006.07103001.x.
McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, *2*, 301–323, http://dx.doi.org/10.1023/A:1014616908334.
Mihalic, S. F., & Irwin, K. (2003). Blueprints for violence prevention. *Youth Violence and Juvenile Justice*, *1*, 307–329, http://dx.doi.org/10.1177/1541204003255841.
Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, *24*, 315–340, http://dx.doi.org/10.1177/109821400302400303.
Muthén, L. K., & Muthén, B. O. (1998–2009). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
Nathanson, L., Sawyer, B., & Rimm-Kaufman, S. E. (2007a). *Classroom Practices Teacher Survey.* : University of Virginia (Unpublished instrument).
Nathanson, L., Sawyer, B., & Rimm-Kaufman, S. E. (2007b). *Classroom practices frequency survey.* : University of Virginia (Unpublished instrument).
National Institute of Child Health and Human Development Early Child Care Research Network (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behavior. *The Elementary School Journal*, *102*, 367–387, http://dx.doi.org/10.1086/499709.
National Institute of Child Health and Human Development Early Child Care Research Network (2005). A day in third grade: A large-scale study of classroom quality and teacher and student behavior. *The Elementary School Journal*, *105*, 305–323, http://dx.doi.org/10.1086/428746.
Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The Journal of Behavioral Health Services and Research*, *39*, 374–396, http://dx.doi.org/10.1007/s11414-012-9295-x.
Northeast Foundation for Children (2007). *The responsive classroom level 1 workbook.* Greenfield, MA: Northeast Foundation for Children.
Northeast Foundation for Children (2012a). FAQ. In responsive classroom. Retrieved June 11, 2012, from. http://www.responsiveclassroom.org/faq
Northeast Foundation for Children (2012b). 2-Page fact sheet. In about responsive classroom. Retrieved October 24, 2012, from. http://www.responsiveclassroom.org/about-responsive-classroom
O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*, 33–84, http://dx.doi.org/10.3102/0034654307313793.
Parcel, G. S., Ross, J. G., Lavin, A. T., Portnoy, B., Nelson, G. D., & Winters, F. (1991). Enhancing implementation of the teenage health teaching modules. *Journal of School Health*, *61*, 35–38, http://dx.doi.org/10.1111/j.1746-1561.1991.tb07857.x.
Payne, A. A., & Eckert, R. (2010). The relative importance of provider, program, school, and community predictors of the implementation quality of school-based prevention programs. *Prevention Science*, *11*, 126–141, http://dx.doi.org/10.1007/s11121-009-0157-6.
Pianta, R. C., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science*, *315*, 1795–1796, http://dx.doi.org/10.1126/science.1139719.
Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008a). Classroom effects on children's achievement trajectories in elementary school. *American Education Research Journal*, *45*, 365–397, http://dx.doi.org/10.3102/0002831207308230.
Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*, 109–119, http://dx.doi.org/10.3102/0013189X09332374.
Pianta, R. C., LaParo, K. M., & Hamre, B. K. (2008b). *Classroom assessment scoring system: manual K-3.* Baltimore, MD: Brookes Publishing.
Pianta, R. C., LaParo, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal*, *102*, 225–238, http://dx.doi.org/10.1086/499701.
Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008c). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *23*, 431–451, http://dx.doi.org/10.1016/j.ecresq.2008.02.001.
Ponitz, C. C., Rimm-Kaufman, S. E., Brock, L. L., & Nathanson, L. (2009a). Early adjustment, gender differences, and classroom organizational climate in first grade. *The Elementary School Journal*, *110*, 142–162, http://dx.doi.org/10.1086/605470.
Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009b). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review*, *38*, 102–120 (Retrieved from http://tinyurl.com/bvbs3uv)
Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review*, *38*, 510–532 (Retrieved from http://tinyurl.com/3ejprn4)
Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in head start settings. *Early Childhood Research Quarterly*, *23*, 10–26, http://dx.doi.org/10.1016/j.ecresq.2007.09.001.
Rimm-Kaufman, S. E., Larsen, R. A., Baroody, A. A., Curby, T. W., Ko, M., Thomas, J. B., et al. (2013). *Efficacy of the responsive classroom approach: Results from a three year, longitudinal randomized controlled trial.* (Manuscript submitted for publication).
Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, *45*, 958–972, http://dx.doi.org/10.1037/a0015861.
Rimm-Kaufman, S. E., Fan, X., Chiu, Y. J., & You, W. (2007). The contribution of the responsive classroom approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology*, *45*, 401–421, http://dx.doi.org/10.1016/j.jsp.2006.10.003.
Rimm-Kaufman, S. E., & Hamre, B. (2010). The role of psychological and developmental science in efforts to improve teacher quality. *Teacher College Record*, *112*, 2988–3023 (Retrieved from http://www.tcrecord.org/library/content.asp?contentid=16046)
Rimm-Kaufman, S. E., & Sawyer, B. E. (2004). Primary-grade teachers' self-efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the Responsive Classroom approach. *The Elementary School Journal*, *104*, 321–341, http://dx.doi.org/10.1086/499756.
Rimm-Kaufman, S. E., Storm, M. D., Sawyer, B. E., Pianta, R. C., & LaParo, K. M. (2006). The teacher belief Q-sort: A measure of teachers' priorities and beliefs in relation to disciplinary practices, teaching practices, and beliefs about children. *Journal of School Psychology*, *44*, 141–165, http://dx.doi.org/10.1016/j.jsp.2006.01.003.
Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., Vincus, A., et al. (2003). Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools. *Health Education & Behavior*, *30*, 375–391, http://dx.doi.org/10.1177/1090198103030003010.
Ringwalt, C., Vincus, A. A., Hanley, S., Ennett, S. T., Bowling, M. J., & Rohrbach, L. A. (2009). The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2005. *Prevention Science*, *10*, 33–40, http://dx.doi.org/10.1007/s11121-008-0112-y.

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors*, *27*, 989–993, http://dx.doi.org/10.1016/S0306-4603(02)00300-3.
Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine*, *22*, 237–260, http://dx.doi.org/10.1006/pmed.1993.1020.
Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation. *Evaluation & the Health Professions*, *29*, 302–333, http://dx.doi.org/10.1177/0163278706290408.
Rohrbach, L. A., Gunning, M., Sun, P., & Sussman, S. (2010). The project towards no drug abuse (TND) dissemination trial: Implementation fidelity and immediate outcomes. *Prevention Science*, *11*, 77–88, http://dx.doi.org/10.1007/s11121-009-0151-z.
Ross, J. G., Luepker, R. V., Nelson, G. D., Saavedra, P., & Hubbard, B. M. (1991). Teenage health teaching modules: Impact of teacher training on implementation and student outcomes. *Journal of School Health*, *61*, 31–34, http://dx.doi.org/10.1111/j.1746-1561.1991.tb07856.x.
Rowan, B., Correnti, R., & Miller, R. J. (2002). What large scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record*, *104*, 1525–1567 (Retrieved from http://tinyurl.com/bocjvd3)
Sawyer, B. E., & Rimm-Kaufman, S. E. (2007). Teacher collaboration in the context of the Responsive Classroom approach. *Teachers and Teaching: Theory and Practice*, *13*, 211–245, http://dx.doi.org/10.1080/13540600701299767.
Selig, J. P., & Preacher, K. J. (2008). Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects [computer software]. Available from. http://quantpsy.org/http://quantpsy.org/
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton-Mifflin.
Smylie, M. A. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. *American Educational Research Journal*, *25*, 1–30, http://dx.doi.org/10.3102/00028312025001001.
Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediated effects of the child development project. *Social Psychology of Education*, *4*, 3–51, http://dx.doi.org/10.1023/A:1009609606692.
Stuhlman, M. W., & Pianta, R. C. (2009). Profiles of educational quality in first grade. *The Elementary School Journal*, *109*, 323–342 (Retrieved from http://tinyurl.com/3ryarw7)
Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child and Family Behavior Therapy*, *24*, 23–50, http://dx.doi.org/10.1300/J019v24n01_03.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.
Wanless, S. B., Patton, C. L., Rimm-Kaufman, S. E., & Deutsch, N. L. (2012). Setting-level influences on implementation of the Responsive Classroom approach. *Prevention Science*, http://dx.doi.org/10.1007/s11121-012-0294-1 (Advance online publication).
Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*, 411–419, http://dx.doi.org/10.1037//0022-0663.89.3.411.
Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*, 287–301, http://dx.doi.org/10.1111/1467.
Wong, H. K., & Wong, R. T. (2009). *The first days of school: How to be an effective teacher.* Mountain View, CA: Harry K. Wong.
Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2007). The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation*, *17*, 191–210, http://dx.doi.org/10.1080/10474410701413145.
Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (Eds.). (2004). *Building academic success on social and emotional learning: What does the research say?*. New York, NY: Teachers College Press.