American Educational Research Journal
June 2014, Vol. 51, No. 3, pp. 567–603
DOI: 10.3102/0002831214523821

Efficacy of the *Responsive Classroom* Approach: Results From a 3-Year, Longitudinal Randomized Controlled Trial

Sara E. Rimm-Kaufman
Ross A. A. Larsen
Alison E. Baroody
University of Virginia
Timothy W. Curby
George Mason University
Michelle Ko
Julia B. Thomas
Eileen G. Merritt
University of Virginia
Tashia Abry
Arizona State University
Jamie DeCoster
University of Virginia

This randomized controlled field trial examined the efficacy of the Responsive Classroom *(*RC*) approach on student achievement. Schools (*n *= 24) were randomized into intervention and control conditions; 2,904 children were studied from end of second to fifth grade. Students at schools assigned to the* RC *condition did not outperform students at schools assigned to the control condition in math or reading achievement. Use of* RC *practices mediated the relation between treatment assignment and improved math and reading achievement. Effect sizes (ES) were calculated as standardized coefficients. ES relations between use of* RC *practices and achievement were .26 for math and .30 for reading. The* RC *practices and math achievement relation was greater for students with low initial math achievement (ES = .89). Results emphasize fidelity of implementation.*

KEYWORDS: *Responsive Classroom*, fidelity of implementation, mathematics, reading, social and emotional learning

Social and emotional learning (SEL) interventions are designed to teach students social and emotional skills considered foundational to academic learning in school and beyond (Durlak, Weissberg, Dymnicki, Taylor, &

Schellinger, 2011). Universal school-based SEL programs that teachers deliver to students in classrooms have been viewed as levers for creating school improvement. SEL skills have been included into state learning standards (Dusenbury, Zadrazil, Mart, & Weissberg, 2011). There have been recent efforts to include the Academic, Social and Emotional Learning Act (HR 1875, 2013) in revisions of the Elementary and Secondary Education Act.

Although research on SEL interventions has accumulated in past decades (e.g., Durlak et al., 2011; Social and Character Development Research Consortium [SACD], 2010), many SEL interventions have not been subject to rigorous efficacy trials examining their impact on student achievement.

Sara E. Rimm-Kaufman is a professor at the University of Virginia in the Curry School of Education and the Center for Advanced Study of Teaching and Learning, 350 Old Ivy Way, Suite 300, Charlottesville, VA 22903; e-mail: *serk@virginia.edu*. She conducts research on elementary school and classroom teaching practices and experiences that support students' engagement in learning, development of self-control, and social and emotional skills.

Ross A. A. Larsen is an assistant professor at Virginia Commonwealth University in the Department of Foundations of Education. His research interests include multilevel clustering, latent growth curve modeling with parallel processes, Monte Carlo simulations, and Bayesian statistics.

Alison E. Baroody is an assistant professor at San Francisco State University in the Department of Child and Adolescent Development. She conducts research on children's interest and engagement in learning, multimethod measurement approaches of assessment, and early childhood and elementary school educational interventions.

Timothy W. Curby is an assistant professor at George Mason University in the Department of Applied Developmental Psychology. His work focuses on the role of teacher-student interactions in promoting young children's development, particularly in relation to their social-emotional development.

Michelle Ko is a research associate at University of Virginia in the Center for Advanced Study of Teaching and Learning. She works on large-scale education research studies as a data manager and data analyst.

Julia B. Thomas is a research associate at the University of Virginia in the Center for Advanced Study of Teaching and Learning. She is the managing director of the University of Virginia Social Development Laboratory where she manages large-scale education research projects.

Eileen G. Merritt is an assistant professor at the University of Virginia in the Curry School of Education and the Center for Advanced Study of Teaching and Learning. She conducts research on teaching practices in science and mathematics classrooms.

Tashia Abry is an assistant research professor at Arizona State University at the T. Denny Sanford School of Social and Family Dynamics. Her research focuses on the implementation of school-based interventions and the role of the classroom ecology in shaping early learning and development.

Jamie DeCoster is a senior research scientist at University of Virginia in the Center for Advanced Study of Teaching and Learning. He works primarily as a statistician and methodologist, and conducts research to determine ways to make the methods used by scientists more accurate, flexible, and efficient.

As a result, policymakers and school decision makers are left with too little information upon which to make decisions. Decision makers question: Will we diminish children's academic achievement if we place increased emphasis on (and allocate more time toward) children's social and emotional learning? This question is worth asking. Classrooms that provide nurturance but do not emphasize academic growth fail to produce student achievement gains (Lee & Smith, 1999; Shouse, 1996). Further, work in developmental psychology calls into question the assumption that improvements in social skills cross over and benefit students' academic skills (Duncan et al., 2007).

The *Responsive Classroom*® (*RC*) approach is a widely used professional development intervention comprised of a set of practical teaching strategies designed to support children's social, academic, and self-regulatory skills. The *RC* approach is designed to enhance teachers' capacity to create a caring, well-managed classroom environment characterized by respectful social interactions and academically engaging instruction (Northeast Foundation for Children, 2007, 2009, 2014a). The *RC* approach was recently endorsed as one of 21 recommended programs by the Collaborative for Academic, Social and Emotional Learning (CASEL, 2013) for its high-quality design, support for implementation, and evidential basis. As a result, school decision makers are particularly interested in the extent to which the *RC* approach relates to academic outcomes.

The present study, Responsive Classroom Efficacy Study (RCES), builds upon existing research (Elliott, 1999; Rimm-Kaufman, Fan, Chiu, & You, 2007) and represents the first randomized controlled trial of the approach. We randomized elementary schools to intervention versus control conditions and followed a cohort of students and their teachers from the end of second grade through fifth grade. Three research questions pertaining to student achievement were addressed:

Research Question 1: What is the impact of the *RC* approach on students' reading and math achievement over 3 years?

Research Question 2: To what extent does fidelity of implementation mediate the relation between treatment assignment (intervention vs. control) and reading and math achievement over 3 years?

Research Question 3: To what extent is the mediational relation affected by whether or not students are qualified for free and reduced priced lunch (FRPL) and by students' initial achievement?

Recent synthesis work on SEL interventions describes the need for randomized controlled trials that examine the impact of SEL interventions on achievement outcomes, attend to issues of fidelity of implementation, and consider the moderators of intervention effectiveness (Durlak et al., 2011; Greenberg, 2010). The present study contributes to a critical area of need.

School-Based Social and Emotional Learning

Randomized controlled trials examining links from SEL interventions to achievement outcomes reveal mixed results. A cluster-randomized controlled trial on *Positive Action*, an intervention with explicit lessons on self-concept, self-management, peer social behaviors, and self-improvement, produced gains in elementary school math and reading (Hedge's *g* effect size [ES] = .50–.72) after 3 years of implementation (Snyder et al., 2009). Findings from the 4Rs ("Reading, Writing, Respect and Resolution"), a program that integrates SEL skills into language arts instruction, demonstrated the presence of gains in reading and math achievement (measured by test scores and teacher report) over 2 years of implementation. Findings were evident only for children whose teachers identified them as showing signs of initial behavioral risk (above a clinical cut-off in conduct problems and aggression) (Jones, Brown, & Aber, 2011). Findings from the Child Development Project, a program designed to promote children's prosocial skills and give children opportunities to experience relatedness, competency, and autonomy, related to enhanced sense of community and improved achievement and behavioral outcomes (Battistich, Schaps, & Wilson, 2004), particularly when schools implemented the intervention as intended (Battistich, Solomon, Kim, Watson, & Schaps, 1995). Research on the Good Behavior Game, a group game designed to decrease disruptive behaviors, combined with an enriched academic curriculum implemented in first grade showed long-term gains in achievement to the end of high school (Bradshaw, Zmuda, Kellam, & Ialongo, 2009). The state of research suggests positive outcomes are more likely in conditions of high implementation and when outcomes align with the theory of change (SACD, 2010). SEL interventions show stronger relations to social than achievement outcomes (Durlak et al., 2011).

Quasi-experimental research links SEL interventions to student achievement. For instance, findings from a nonrandomized field trial of a multicomponent intervention administered to fifth and sixth graders designed to foster bonding toward school was associated with math and reading achievement gains (Abbott et al., 1998). A quasi-experimental study of the RULER Feelings Words Curriculum, an intervention designed to teach children how to interpret and communicate emotions, demonstrated that fifth and sixth graders in classrooms using the intervention outperformed students in the control group on language arts grades, albeit a small effect (Brackett, Rivers, Reyes, & Salovey, 2010). A meta-analysis reviewing 213 SEL interventions (including a range of designs) showed a positive contribution of SEL programming on student behavioral and emotional skills as well as academic performance. Effect sizes (Hedge's *g*) were .57 for gains in social and emotional skills and .27 for gains in achievement (Durlak et al., 2011).

The *Responsive Classroom* Approach

The *RC* approach was developed by the Northeast Foundation for Children (NEFC); more than 120,000 teachers have been trained in the approach. According to NEFC (2014a), the *RC* approach is an "approach to elementary education that gives teachers the tools they need to be highly effective instructors" (p. 2). Like many professional development interventions, the *RC* approach continues to develop to match contemporary needs. In this paper, we describe the *RC* approach based on the version available in 2008-2011. The *RC* approach offers principles including: "The social curriculum is as important as the academic curriculum; How children learn is as important as what they learn: Process and content go hand in hand; Knowing the children we teach—individually, culturally, and developmentally—is as important as knowing the content we teach" (NEFC, 2014b, "Guiding Principles"). Ten practices emanate from the principles. Teachers and children gather together daily for a Morning Meeting to create a sense of community among teachers and students and instill curiosity about the learning ahead. Teachers use Academic Choice to structure lessons to provide students autonomy and choice. The practices are described briefly in the appendix in the online journal and extensively in manuals and books (NEFC, 2007, 2009). Training in the *RC* approach involves two 1-week workshops, each lasting 35 hours, typically taken in consecutive summers. In-school coaching follows each workshop.

The *RC* approach differs from prevalent approaches to SEL. For example, the *RC* approach emphasizes *how* to teach rather than *what* to teach. Instead of establishing a set curriculum for teaching SEL skills (i.e., those in RULER or Promoting Alternative Thinking Strategies [PATHS]), the *RC* approach embeds modeling of prosocial behavior, collaboration, and self-control into instructional practices. *RC* practices are designed to align with existing curricula in the school rather than introducing content with an SEL focus (e.g., in 4Rs).

Figure 1 presents a logic model that describes a theory of change for the *RC* approach (Knowlton & Phillips, 2008). In theory, training and coaching in the *RC* approach leads to teacher change (use of *RC* practices), which leads to enhanced emotional support, proactive classroom management, and in turn, student motivation and engagement followed by improved student achievement. Two hallmark characteristics stand out. First, *RC* practices are designed to create emotionally supportive classroom relationships. *RC* practices are intended to improve teacher-student and peer relationships within classrooms so children develop prosocial skills, cooperate with one another, and perceive that their teachers and peers care about them. Second, *RC* practices offer a proactive approach to classroom management. Using *RC* classroom practices *early in the year* and *early in each day* is key to their use. Teachers set expectations for behavior and learning with the

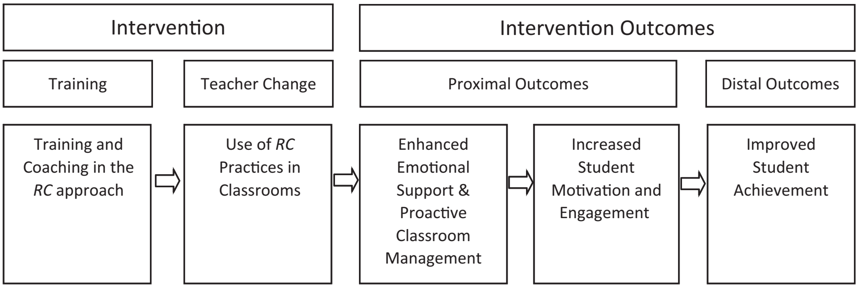

Figure 1. **Logic model describing the theory of change for the *Responsive Classroom* approach.**

goal that children will internalize expectations over time, develop self-regulatory skills, and behave more autonomously.

Small-scale studies suggest the positive contribution of the *RC* approach. Research examining schoolwide implementation of the *RC* approach using a pre-post design showed gains in social skills (d = .34–.41) (Elliott, 1999). A quasi-experimental study compared students at intervention versus control schools. Teachers' use of *RC* practices related positively to teachers' perception of children's reading achievement, teachers' perception of closeness toward their students, children's assertiveness, and children's prosocial behavior (Rimm-Kaufman & Chiu 2007). Analyses compared student test scores of children at *RC* versus control schools for three cohorts of elementary school students exposed to the *RC* approach for 1, 2, and 3 years. Findings showed that the *RC* approach contributed to the gains in both reading and math after controlling for poverty and previous years' test scores (d = .16–.21 for reading; d = .16–.39 for math; Rimm-Kaufman et al., 2007). The magnitude of the findings is comparable to those found for SEL interventions in meta-analytic work (Durlak et al., 2011). The state of existing work indicates the need for a randomized controlled trial to assess efficacy. Thus, the first research question addressed in the present study asks: What is the impact of 3 years of the *RC* approach on student achievement? We hypothesize higher achievement among children in intervention than control schools, demonstrating the positive impact of the *RC* approach.

Fidelity of Implementation

The *RC* approach is a fully developed, manualized intervention with a well-documented training process (NEFC, 2007, 2009). Despite efforts to create consistency among teachers using the *RC* approach, existing work on fidelity of implementation on the *RC* approach (Abry, Rimm-Kaufman, Larsen, & Brewer, 2013) as well as other SEL interventions (Reyes, Brackett, Rivers, Elbertson, & Salovey, 2012) suggests that uptake of the intervention varies across schools and teachers. The process of adopting SEL interventions involves a complex process as teachers change their practices and beliefs to align with the intervention (Evans, 2001; Rimm-Kaufman, Storm, Sawyer, Pianta, & La Paro, 2006). Teachers may receive training but not use the *RC* practices, adopt some *RC* practices but not others, or adopt superficial elements of *RC* practices but disregard underlying principles. Because the *RC* logic model rests upon the assumption that teachers implement *RC* practices with fidelity, fidelity of implementation needs to be measured carefully. Century (2010) defines fidelity of implementation as "the extent to which the critical components of an intended program are present when that program is enacted" (p. 202). We operationalize fidelity of implementation to the *RC* approach as frequency and adherence to the 10 critical *RC* practices (e.g., Morning Meeting, Academic Choice, Interactive Modeling, etc. as described in the appendix in the online journal).

Two challenges stand out in relation to measuring fidelity in the *RC* approach. First, the *RC* approach is designed to modify *how* teachers interact with students in their classrooms instead of *what* students are learning. Many SEL interventions are comprised of a sequenced curriculum (CASEL, 2013), and measuring fidelity involves assessing the presence/absence of curricular activities and time spent delivering content (Bickman et al., 2009). In contrast, measuring fidelity of the *RC* approach requires assessment of the frequency of use of *RC* practices (e.g., Do teachers use Morning Meeting once a day, once a week, or once a month?) as well as the extent to which implementation adheres to *RC* principles and practices (e.g., Do teachers provide students choices about what kind of work to do and/or how to do the work?)

The *RC* approach presents yet another challenge in measuring fidelity (Cordray & Pion, 2006). *RC* practices were derived from well-known educational and developmental theories (e.g., Piaget, Vygotsky, Dewey), and therefore, some *RC* practices resemble teaching practices used in classrooms by teachers who have not been trained in the *RC* approach. Further, the intervention is described and characterized in a series of books published and distributed widely by NEFC, resulting in widespread adoption of intervention components (e.g., Denton, 2005; Denton & Kriete, 2000; Kriete, 2002). *RC* practices share common qualities with what is considered typical elementary school practices. Many teachers use a class meeting to start the school day, and features of those class meetings (e.g., student sharing, fun group activities) resemble the *RC* Morning Meeting. Many teachers incorporate differentiated instruction into their classrooms, a practice that shares common elements with Academic Choice. To address this issue, we assessed use of *RC* practices in both intervention and control schools using measures free from *RC* terminology. The measures of *RC* practices reflect fidelity of implementation in intervention schools and use of practices that resemble *RC* practices in the control condition. (We use the terms *fidelity* and *use of RC practices* interchangeably.)

Thus, the second research question addressed examines fidelity of implementation. To what degree does teachers' fidelity of implementation (in third, fourth, and fifth grade) mediate the relation between assignment to the *RC* condition and fifth-grade achievement? We hypothesized that teachers' fidelity of implementation will mediate the relation between the *RC* approach and reading and mathematics achievement.

Student Subgroups

SEL interventions may not relate equivalently to outcomes for all students. Some students come to school with more advanced academic skills than others (Li-Grining, Votruba-Drzal, Maldonado-Carreño, & Haas, 2010). To understand the extent to which findings about teachers' use of *RC* practices can be generalized, research needs to examine *for whom* the intervention is effective (Bloom & Michalopoulos, 2010). The present study examines potential moderated mediation effects,[1] testing the extent to which use of *RC* practices mediates the relation between *RC* training and achievement more for some children compared to others. Poverty (eligibility for FRPL) and low initial achievement (performance below 25th percentile in math achievement at the end of second grade) were potential moderators.

Students living in poverty tend to show lower achievement than their non-poor counterparts, an association that has been attributed to less predictable living situations and reduced access to social, material, and learning resources at home (Gershoff, Aber, Raver, & Lennon, 2007; Yoshikawa, Aber, & Beardslee, 2012). Links between poverty and lower student achievement raise questions about the degree to which supportive classroom experiences can partially ameliorate risk associated with poverty. Emotionally and instructionally supportive school experiences appear to be particularly important for children from poor families (Crosnoe et al., 2010; Hughes & Kwok, 2007).

Low-achieving children are the focus of national efforts to decrease the achievement gap, and there is a critical need for improved understanding of the instructional practices that support learning for children low in math achievement (National Mathematics Advisory Panel, 2008). Although low achievement in early elementary school tends to persist into the middle and upper elementary school grades (Claessens, Duncan, & Engel, 2009), students' course of achievement is not firmly established and classroom experiences can shift students toward more positive growth trajectories. Exposure to high-quality classroom environments characterized by social and emotional supportiveness (Pianta, Belsky, Vandergrift, Houts, & Morrison, 2008), nonconflictual relationships with teachers, and instruction requiring analysis and inference partially offsets the effects of initial low achievement (Crosnoe et al., 2010).

Efforts to improve schooling for poor and low-achieving children demand research that examines whether teaching practices such as *RC* practices are more efficacious for some children than others. The third research question asks: To what extent is the mediated relation (i.e., *RC* condition to fifth-grade achievement via fidelity of implementation) moderated by child characteristics? We hypothesize that the mediated path will be stronger for children who qualify for FRPL and children with low initial achievement.

Other Contributing Factors

Existing research directs attention to other attributes of students, teachers, and schools that need to be considered as potential covariates in predicting achievement. Student gender is often linked to math and reading achievement (Robinson & Lubienski, 2011). Student poverty (Votruba-Drzal, 2006) and English language learner (ELL) status (Kieffer, 2008) needed to be considered given their relation to achievement. We also included test form administered (plain English versus standard) as a potential covariate. Student initial achievement was an important covariate to include because of the relative stability in achievement from third to fifth grade (Claessens et al., 2009); math achievement at the end of second grade served as a proxy for initial student achievement.

There are teacher- and school-level features that contribute to student achievement (Bradshaw, Mitchell, & Leaf, 2010) and warrant consideration as covariates. Teaching draws upon teachers' experiences and educational background; teachers' years of experience and educational attainment were selected to represent the presence of these resources (Crosnoe & Cooper, 2010). Further, school-level poverty and student achievement (aggregated to the school level) were included because of theory and empirical work linking these variables to student performance (Conduct Problems Prevention Research Group, 2010; Raver, 2012).

Method

Participants

This study included 24 elementary schools in a large ethnically and socioeconomically diverse school district located in a mid-Atlantic state. District policy required all elementary schools to select an approach to foster SEL or behavioral learning, a policy implemented over a 3-year period. When the district was invited to participate, 24 elementary schools had not made selections and/or initiated formal training and were willing to participate in a randomized controlled trial of the *RC* approach. Centralized district administrators collaborated with the research team to invite school principals to participate in the study; 100% of the schools invited agreed to participate.

Schools were assigned randomly to intervention or waitlist control condition. The methodologist generated random numbers for each of the 24 schools using the random number function in Excel. The random number values assigned to the schools were sorted; the top 12 schools were assigned to the treatment condition and the bottom 12 schools were assigned to the control condition. Tests were conducted to ensure treatment and control schools were comparable (on gender, FRPL, ethnic composition, ELL status). After random assignment, it became apparent from an informational survey that 2 schools in the control condition had received low-level exposure to the *RC* intervention. In each school, the principal reported that one of the third-, fourth-, or fifth-grade teachers (<15% of sampled teachers at each school) had received a 1-week *RC* training. For example, at one school, the principal and some teachers acquired two NEFC-published books and used them to self-teach *RC* practices. Because of the low-level contamination, one of the two schools was selected at random and placed into the intervention condition, resulting in 13 intervention and 11 control schools. Control schools were exempt from the district policy requiring selection of an SEL or behavioral learning approach for the study duration. Student initial math achievement (Stanford 10) was gathered after randomization. There were no statistically significant differences between treatment and control schools on gender, FRPL, ethnic composition, ELL status, and student initial math achievement based on independent t tests and logistic regressions at baseline. See Table 1.

Teacher participants (n = 276) were third-, fourth-, and fifth-grade teachers during the years of 2008–2009, 2009–2010, and 2010–2011, respectively. Teachers (95 third, 92 fourth, and 89 fifth) were primarily female (n = 248; 90%) and had, on average, 10 years of experience (range, 1–38). Teachers were mostly Caucasian (n = 232; 84%) with other ethnic groups represented (11 [4%] African American, 6 [2%] Asian, 10 [4%] Hispanic, and 17 [6%] Other).

All second graders attending the 24 study schools in spring 2008 were enrolled in the study. Students entering third grade in 2008–2009, fourth grade in 2009–2010, and fifth grade in 2010–2011 at the 24 schools were included as participants. Two rationales guided student participant decisions: (1) Randomization occurred at the school level, and thus, our model included typical patterns of school transience (including exit and entrance) to improve ecological validity; (2) school-level clustering reduced statistical power, and to compensate, full information maximum likelihood (FIML) techniques were used to handle missing data. (Additional information about missing data, school-level sampling, and analysis can be found in the analytic approach section.) In the final sample of students (n = 2,904) 1,422 (49%) were female and 929 (32%) were eligible for FRPL. The sample was ethnically diverse (1,191 [41%] Caucasian, 319 [11%] African American, 552 [19%] Asian, 697 [24%] Hispanic, 145 [5%] other). Schools identified 813 (28%) as ELL. See Table 1 for description by treatment condition.

Figure 2 describes the pattern of students' entrance and attrition from the sample from 2008 to 2011. The magnitude of student attrition (intervention, n = 395; control, n = 350) and student entrance (intervention, n = 527; control, n = 444) was comparable ($p > .05$) between intervention and control conditions between end of second and fifth grades. Students (n = 109) in French and

Table 1
Descriptive Statistics for Variables by Treatment and Control Condition (n = 24)

Variable	Treatment (n = 13)			Control (n = 11)			
	%	*M*	*SD*	%	*M*	*SD*	*t*
School level							
Years of teaching experience		10.40	3.81		11.75	4.02	0.84
Teachers with MS degrees	67			65			0.68
Percentage female	51			46			1.69
Free/reduced priced lunch	40			29			1.09
Caucasian	37			45			–1.14
African American	12			10			0.48
Asian	18			18			–0.02
Hispanic	28			22			1.02
Other	5			5			–0.33
Percentage ELL	32			27			0.69
Initial math achievement (Stanford 10)		578.06	14.46		583.69	17.06	–0.88
Child level	Treatment (n = 1,467)			Control (n = 1,437)			
Initial math achievement (Stanford 10)		579.31	46.03		584.63	42.00	
Female	51			47			
Free/reduced priced lunch	38			27			
Caucasian	38			45			
African American	11			10			
Asian	18			20			
Hispanic	28			20			
Other	5			5			
ELL status		0.99	1.51		0.79	1.35	
Plain English test	2			4			
Fifth-grade math		526.31	74.33		534.45	67.99	
Fifth-grade reading		492.30	64.84		501.21	62.71	

Note. ELL = English language learners.

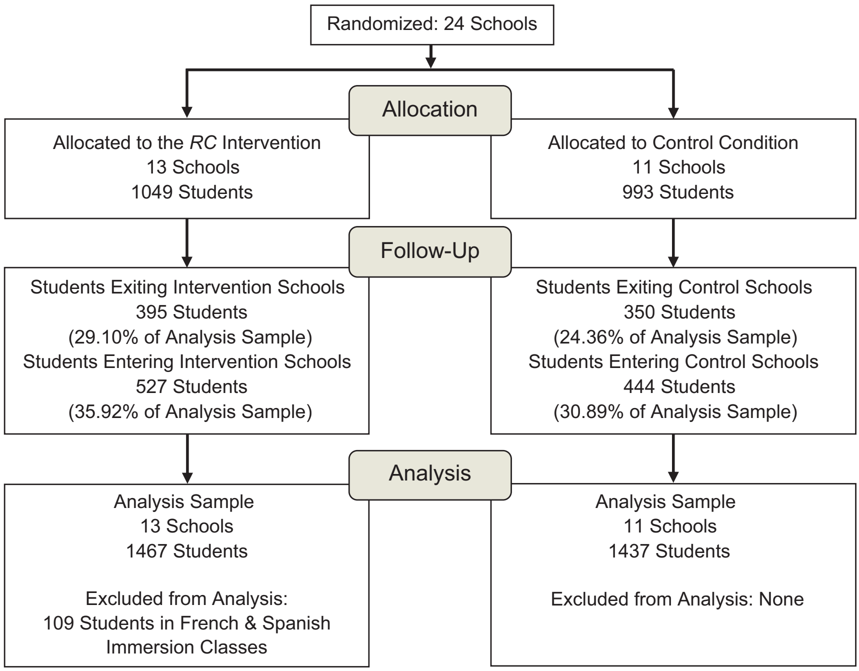

Figure 2. **Research design and attrition/addition flow through the *Responsive Classroom* Efficacy Study.**

Note. All students were included in the final analysis in keeping with the school-level analysis.

Spanish immersion classrooms were present only in the intervention group and were excluded from analyses. Logistic regressions conducted to examine differential entrance and attrition between intervention and control conditions by gender, FRPL, ethnicity, and ELL showed no statistically significant differences between intervention and control schools ($p > .10$).

Intervention

Third-, fourth-, and fifth-grade teachers in the intervention condition received training in the *RC* approach, involving two 1-week long training sessions, *RC* 1 and *RC* 2, delivered in two consecutive summers. Third- and fourth-grade teachers received *RC* 1 in summer 2008 and *RC* 2 in summer 2009. Fifth-grade teachers received *RC* 1 in summer 2009 and *RC* 2 in summer 2010. Teachers received 3 days of school-based coaching and opportunities to engage in three *RC* workshops during each school year subsequent to training. Further, each teacher received *RC* manuals, two additional *RC* books, on-demand support (in person and via e-mail) from coaches, and articles on the *RC* website. School administrators at the 13 schools received *RC* 1 and *RC* 2 training and three sessions per year of administrator coaching led by NEFC consultants. In addition, NEFC consultants met with each school administrator for a planning meeting in fall and spring. Each school received a full set of *RC* books for their library. All training and coaching sessions were conducted by NEFC consultants who were trained and certified *RC* instructors.

In *RC* 1, participants learned strategies to implement key *RC* practices, including Morning Meeting, Rule Creation, Interactive Modeling, Positive Teacher Language, and Logical Consequences. Additional *RC* practices, Guided Discovery, Academic Choice, Classroom Organization, and Collaborative Problem-Solving, were emphasized in *RC* 2.

Teachers in the control group did not receive *RC* training and continued with "business as usual" approaches. Principal interviews and questionnaires were used to assess the social and emotional learning and classroom management practices among schools in the control condition. No principals reported use of the *RC* approach at the school level. Seven principals reported no schoolwide program, two reported schoolwide reward systems, one described schoolwide teams using the Kagan approach, and one described use of Positive Behavioral Support.

Procedures

Schools were recruited in November 2007 and randomized into condition. Baseline data collection was conducted in spring 2008 by administering the Stanford 10 mathematics test (described below) to all second graders at participating schools.

Classroom observations and teacher questionnaires to measure fidelity were conducted in 2008–2009 for third-grade teachers, 2009–2010 for fourth-grade teachers, and 2010–2011 for fifth-grade teachers. Thus, study year data were collected on third-grade teachers in the school year immediately following *RC* 1 training corresponding to their first year of implementation. Study year data were collected on fourth- and fifth-grade teachers in the year subsequent to *RC* 2 training in their second year of implementation. This decision was based on grant time constraints, but is consistent with existing patterns in *RC* training. NEFC states that more than one-half of teachers (nationally) who receive training in the *RC* approach attend *RC* 1 only.

Five classroom observations were conducted throughout the school year for each teacher to measure observed fidelity of implementation. For each teacher, two observations were conducted during the first hour in the morning and three observations were conducted for one hour during math lessons. Observations were scheduled following a set of decision rules to ensure evenness over the year and across schools and assignment condition.

Observations corresponded to three windows (window 1: late September to late November; window 2: late November to mid February; and window 3: late February to late April). Morning observations were conducted in two of the three windows with no more than two-thirds of the teachers at each school receiving a morning observation per window. Math observations were conducted in each of the three windows. Research assistants conducted videotaping (for work beyond the scope of this article) and rated fidelity of implementation using the Classroom Practices Observational Measure (described below) for the full 60-minute observations. In April of each year, teachers were surveyed to gather information on teachers' years of experience, teachers' level of education, and two teacher-report measures of fidelity of implementation, the Classroom Practices Teacher Survey and Classroom Practices Frequency Survey (described below). Teachers received $100 upon completion.

In May 2011, all students present in the fifth grade at the 24 schools were given the fifth-grade state standardized test, the Standards of Learning (Virginia Department of Education [VDOE], 2010). Students who were not English proficient took the plain English math assessment, an equivalent, alternative form of the Standards of Learning math test. Student demographic and standardized test score data were gathered from school records.

Measures

School assignment to treatment condition (intervention vs. control) was the key independent variable. Fidelity, aggregated to the school level, was assessed as a potential mediator. Student math and reading achievement in fifth grade were assessed as outcome variables. Students' FRPL status and initial achievement were included as focal variables in analyses testing for moderated mediation. Child-level covariates included student demographic characteristics (gender, FRPL, ethnicity, ELL), test form (plain English vs. standard), and initial achievement. Four variables were aggregated to the school level and treated as covariates: teacher years of experience, teacher education, student FRPL, and initial math achievement.

Fidelity of Implementation

Observed fidelity of implementation. The Classroom Practices Observation Measure (CPOM) is a manualized observational measure of teachers' use of *RC* practices (fidelity) developed by the research team with feedback from NEFC (Abry, Brewer, Nathanson, Sawyer, & Rimm-Kaufman, 2010). Raters observed and assessed teachers' use of *RC practices* during a 60-minute period. Teachers were rated on 16 items during the morning observation (α = .87) and a subset of 10 items during math instruction (α = .65). (The six Morning Meeting items were relevant to morning observations only.) Each item was written without *RC* language so items could be rated by research assistants with no knowledge of the *RC* approach and could be used in control and intervention classrooms. Internal consistency was comparable in intervention and control conditions. CPOM was coded live, on site using the 3-point Likert scale (1 = *not at all characteristic*, 2 = *moderately characteristic*, 3 = *very characteristic*). The manual includes descriptions and examples of practices that exemplify values of 1, 2, or 3 for each item. One example item is, "Three to five general, positively worded rules are posted in the classroom." The observer rating was based on three criteria: (a) number of rules (3 to 5 vs. less or more), (b) generalness of rules ("Be respectful" not "No talking during silent reading"), and (c) rule positivity ("Take turns" not "Don't yell"). The observer coded 1 if there were no rules posted or if posted rules met one of the three criteria; 2 if two of the three criteria were met; and 3 if all criteria were met. Other items were: "Teacher asks questions or makes statements that invite students to remember expected behaviors" and "Students make individualized choices related to an academic lesson or goal." CPOM scores were computed as the mean of all items across five observations.

Two-day CPOM trainings were conducted by research assistants following a written protocol. Initial reliability was established using eight 60-minute master-coded videos, at minimum. Each coder exceeded 80% exact agreement on eight 1-hour segments prior to initiating data collection (values were 84%, 80%, and 84% exact match for year 1, 2, and 3, respectively). Ongoing monthly drift meetings were conducted for calibration. Each coder viewed and individually coded a 60-minute classroom observation video prior to the meeting. Codes brought to meetings were used to calculate percentage agreement, resulting in an average of 85% exact match between coders and master coders over 3 years.

Teacher reported fidelity of implementation. Teachers reported on their fidelity of implementation using two measures. The Classroom Practices Teacher Survey (CPTS; Nathanson, Sawyer, & Rimm-Kaufman, 2007a) is a 46-item measure assessing adherence to *RC* practices (α = .91). CPTS items did not contain any *RC* terminology, and internal consistency was comparable in intervention and control conditions. Teachers rated each item on a 5-point Likert scale (1 = *not at all characteristic*, 2 = *a little bit characteristic*, 3 = *moderately characteristic*, 4 = *very characteristic*, 5 = *extremely characteristic*). Examples include: "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." CPTS composite scores were computed as means of the 46 items.

The Classroom Practices Frequency Scale (CPFS; Nathanson et al., 2007b) is an 11-item measure assessing teachers' perceived frequency of use of *RC* practices over the course of the school year (α = .88). Internal consistency was comparable in intervention and control conditions. Items were rated on an 8-point scale ranging from 1 = *almost never used the practice*, 2 = *1 × per month*, 3 = *2–4 × per month*, 4 = *1 × per week*, 5 = *2–3 × per week*, 6 = *4 × per week*, 7 = *1 × per day*, 8 = *used the practice more than once a day*. Items did not use *RC* terminology. Items include: "When a rule or procedure is introduced, I asked 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." Mean values were computed for the 11 items.

Fidelity factor score. Confirmatory factor analysis was used to combine the three fidelity measures (CPOM, CPTS, and CPFS) into a single fidelity variable that weighted each measure differentially according to its contribution to a single underlying latent variable. Resulting factor loadings exceeded .90. Cronbach's alpha for the factor score equaled .94 and indicated unidimensionality. Factor scores were used in subsequent fidelity analyses.

Student Achievement Outcomes

Fifth-grade achievement. The paper version of the state standardized test, the Standard of Learning (SOL), was used to assess mathematics and reading achievement in May of students' fifth-grade year (VDOE, 2008). The mathematics test comprised of 50 multiple choice items (α = .88) tapping students' procedural knowledge and conceptual understanding of four skill categories: (a) number and number sense, (b) computation and estimation, (c) measurement and geometry, and (d) probability, statistics, patterns, functions, and algebra (VDOE, 2010). The reading test is comprised of 40 multiple choice items (α = .85–.87) to measure two skills: (a) students' reading and print material comprehension and (b) use of word analysis strategies and information resources. A state data team summed the number of correct items, converted the value to a scaled score (ranging from 0 to 600), and transmitted data to the district. The research team garnered scores from the district.

VDOE (2008) describes test development, calibration, and validity. Test items were developed through collaboration among Virginia educators, VDOE, Educational Testing Service, Pearson, and content experts based upon test blueprints. Calibration was established using Rasch modeling and the Partial Credit model. Test validity was established by gathering empirical evidence supporting the face validity, intrinsic rational validity, content validity, and construct validity. VDOE (2008) offers the standard and plain English versions for math.

Covariates and Variables for Moderated Mediation

Student characteristics. Baseline student math achievement was gathered using the Stanford 10 abbreviated mathematics test (Harcourt Educational Measurement, 2003). Gender, FRPL, ethnicity, ELL status, and test form administered (regularly or plain English version of the math SOL test) were gathered from district data. Gender was coded as 1 = *female*, 0 = *male*. Qualification for FRPL (1 = *yes*, 0 = *no*) was defined as $40,793 for a family of four, roughly below 180% of the federal poverty guideline. Ethnicity was based upon parent report of ethnicity upon student matriculation. ELL status was based on a district-developed scale and cutpoint: students were assigned a value of 1 if they were receiving ELL services and a value of 0 if they were not receiving ELL services.

Students completed the plain English math version of the mathematics test if the student's teacher designated that the student had a learning disability or limited English proficiency. The plain English math version used the same scoring system as the SOL and was designed to be equivalent to standard math assessment (VDOE, 2008). Test form data were garnered from the district and coded 1 = *plain English*, 0 = *standard form*.

Baseline student math achievement, gender, FRPL, ethnicity, ELL, and test form were included as student covariates. FRPL and low initial math achievement were used in moderated mediation analyses. The cutoff for low initial math achievement was below the 25th percentile on the Stanford 10, a decision reflecting review of comparable decisions (Baker, Gersten, & Lee, 2002).

School-level covariates. Data on teacher characteristics (years of teaching experience, level of education) were gathered in a teacher survey. Years of teaching experience referred to total years of teaching experience; level of education was coded as 1 = *master's degrees or higher*. Years teaching experience and level of education were averaged across study years and aggregated to the school level for analysis. School characteristics were gathered from school record data and included gender, percentage of students eligible for FRPL, school ethnic composition, and percentage of ELL students. School-level achievement at baseline was based upon the Stanford 10 mathematics test aggregated up to the school level.

Analytic Approach

Structural equation modeling (SEM) analyses were conducted to test the (a) main effect of treatment on outcomes, (b) the mediated effect of treatment on outcomes through fidelity, and (c) moderated mediation, examining the extent to which the mediated effect of treatment on outcomes through fidelity varies with students' eligibility for FRPL and initial ability level. Analyses were conducted using Mplus 6.12 (Muthén & Muthén, 1998–2010) using a maximum likelihood estimator with TYPE=COMPLEX to control for clustering. Clustering was conducted at the school level, corresponding to the level of randomization and consistent with Murray, Hannan, and Baker (1996). A Monte Carlo integration estimator was used to achieve convergence. (Note that this estimator does not produce traditional fit statistics.) Initially, math and reading achievement outcomes were included in the same model but the model did not converge. Therefore, separate models were conducted for each outcome. Ethnicity variables were included in SEM analyses but results are not reported in accordance with the district memorandum of agreement.

The decision to analyze fidelity of implementation data at the school level instead of the teacher level requires further explanation. Data were gathered on students over 3 years. The classroom composition changed over the 3 years; students nested in one classroom in third grade were often scattered among various classrooms in fourth grade and a different set of classrooms in fifth grade. Such nesting is difficult to model. In addition, some of the schools had students move among classrooms for reading, math, and other subjects, and therefore, students were exposed to two, three, or four teachers on any given day. For example, students may have been exposed to low level of *RC* practices in the morning in one classroom but a higher level of *RC* practices later in the day in another classroom. We were unable to gather data on which children were exposed to which teacher and for how long within each year. Thus, we aggregated fidelity to school level because it matched the level of randomization and addressed changes in the nesting structure across the 3 years and across typical schools days. Findings can be interpreted in relation to what schools comparable to those sampled could expect from adopting and implementing the *RC* approach over a 3-year period.

SEM assumes normality of the endogenous variables, linear relations between variables, independence of observations, and independent exogenous variables. Data were examined using histograms, bivariate correlation tables, QQ-plots, variance inflation factors, and other methods. The data met assumptions of normality, linearity, and independent exogenous variables.

As depicted in Figure 2, students entered and exited the 24 study schools over the 3 years. Student attrition from third to fifth grade was considerable with missing data of approximately 30% for both the fifth-grade math and reading achievement. As mentioned previously, there was no relation between any of the covariates/outcomes and missing data at the school level. Thus, the data were considered to be missing at random (MAR). Traditional missing data techniques such as listwise deletion, pairwise deletion, or mean imputation have been shown to have lower power and more bias of the parameters in the presence of MAR data (Allison, 2002; Enders & Bandalos, 2001; Schafer & Graham, 2002). Thus, missing data in this model were handled through FIML. FIML uses all available data to estimate parameters and thus leads to unbiased parameter and standard error estimates (Little & Rubin, 1987).

We observed substantial ceiling effects for the outcomes; 23.7% of the students scored the maximum math score, 5.9% scored the maximum reading score. Thus, outcomes showing ceiling effects were treated as censored variables (Klein & Moeschberger, 1997), a process involving the treatment of maximum scores as lower bounds for students' true latent score. This approach to accounting for censored outcomes leads to less biased estimates (Long, 1997).

Standardized coefficients were calculated based on the unstandardized coefficients emanating from MPlus output in order to present a coefficient indicating effect size. In the case of a continuous covariate the following formula was used:

$$\beta = (S_X B_{raw})/S_y$$

S_X is the estimated standard deviation (*SD*) of the continuous covariate, B_{raw} is the raw coefficient produced by Mplus, and S_y is the *SD* of the predicted variable. The standardized coefficient is interpreted as the expected number of *SD*s that *y* (the outcome) would increase with a one *SD* increase in *x* (the continuous predictor). The following formula was used for categorical covariates:

$$\beta = B_{raw}/S_y$$

The second formula lacks S_X in the numerator because categorical covariates do not have meaningful *SD*s. The standardized coefficient in this case can be interpreted as the expected number of *SD*s that *y* (the outcome) would increase as *x* (the categorical predictor) increases from 0 to 1 (i.e., from control to treatment condition). Standardized coefficients for indirect effects were calculated by multiplying standardized coefficient of the treatment to fidelity path by the standardized coefficient of the fidelity to outcome path.

Results

Descriptive Statistics

Table 1 presents the means and *SD*s for all independent and dependent variables and covariates for intervention and control schools at the school and child levels. Table 2 shows school-level correlation coefficients. The correlation between treatment assignment and the fidelity of implementation factor was very high ($r = .91$, $p < .01$). Fifth-grade reading and math scores were correlated ($r = .53$, $p < .01$). Initial achievement levels (aggregated to the school level) correlated positively with fifth-grade achievement for reading ($r = .78$, $p < .01$) and math ($r = .50$, $p < .05$). Higher concentration of students at the school eligible for FRPL related to lower initial achievement ($r = -.75$, $p < .01$), fifth-grade math ($r = -.65$, $p < .01$), and reading achievement ($r = -.89$, $p < .01$).

Intraclass correlation (ICC) values were computed to determine the percentage of school-level versus child-level variance for fifth-grade math and reading achievement. ICC values were 0.04 and 0.14, indicating that 4% of the total math variance and 14% of the total reading variance could be attributed to the school level. Thus, analyses accounted for nested effects.

Table 2
Correlation Coefficients Between Pairs of School-Level Variables (*n* = 24)

Variable	Fifth-Grade Math	(A)	(B)	(C)	(D)	(E)	(F)
Fifth-grade reading (A)	0.53**						
Treatment (B)	–0.27	–0.17					
Fidelity (C)	–0.25	–0.21	0.91**				
Years experience (D)	0.32	0.09	–0.18	–0.29			
%-age master's degrees (E)	–0.38	–0.22	0.11	0.14	–0.24		
School-level FRPL (F)	–0.65**	–0.89**	0.19	0.35	–0.14	0.27	
School-level initial math achievement (Stanford 10)	0.50*	0.78**	–0.18	–0.29	0.14	–0.20	–0.75**

Note: FRPL = free/reduced-price lunch.
$^{*}p < .05$. $^{**}p < .01$.

Table 3
School-Level Descriptive Statistics and Differences in Fidelity of Implementation Between Intervention and Control Conditions

	Intervention		Control			
Data source	*M*	*SD*	*M*	*SD*	*t*	Cohen's *d*
Observed fidelity (CPOM)	1.74	0.09	1.30	0.14	9.30***	3.18
Teacher-report adherence (CPTS)	3.99	0.24	3.18	0.27	7.78***	3.19
Teacher-report frequency (CPFS)	2.33	0.36	1.57	0.23	6.03***	2.47
Fidelity of implementation composite	0.59	0.35	–0.71	0.32	9.41***	3.88

$^{***}p < .001$.

Intervention and control schools differed in fidelity of implementation ($p < .001$), as evidenced by *t* tests. See Table 3. Differences were large; effect size Cohen's *d* values ranged from 2.47 to 3.88.

Research Question 1

The results for the two separate SEMs testing the effect of treatment on fifth-grade math and reading achievement are found in Table 4. Treatment was not significantly related to fifth-grade math or reading achievement ($\beta = –.13$ for math, $\beta = –.06$ for reading). Girls had significantly higher math and reading achievement scores than boys. FRPL eligibility was associated negatively to math and reading achievement. Both ELL status and completion of the plain English test related negatively to math and reading. Initial math achievement related positively to fifth-grade math and reading. The

Table 4
Main Effect Model of Treatment on Fifth-Grade Math and Reading Achievement

	Math Achievement			Reading Achievement		
Variable Name	*B*	*SE*	β	*B*	*SE*	β
Child level						
Gender (female)	11.63**	4.43	0.16	14.24**	2.50	0.22
Free/reduced price lunch	–11.16*	4.37	–0.16	–11.96**	3.80	–0.19
ELL status	–5.07*	2.21	–0.10	–7.57**	1.59	–0.17
Plain English test	–40.21**	11.75	–0.56	–28.81**	9.75	–0.45
Initial math achievement	1.10**	0.11	0.63	0.76**	0.05	0.49
School level						
Treatment status	–9.48	6.19	–0.13	–3.95	4.77	–0.06
Years of experience	0.79	1.02	0.04	–0.32	0.55	–0.02
Percentage of master's degrees	–14.06	21.90	–0.03	–6.59	16.78	–0.01
Percentage FRPL (school level)	–0.20	0.18	–0.07	–0.23	0.15	–0.08
Initial math achievement (school level)	–0.26	0.32	–0.06	0.70**	0.24	0.17

Note. R^2 = 0.35 for math achievement. R^2 = 0.45 for reading achievement. Child-level initial math achievement is group-mean centered. ELL = English language learner.
$^*p < .05$. $^{**}p < .01$.

models accounted for a substantial portion of the variance in achievement outcomes (R^2 = .35 for math, R^2 =.45 for reading).

Research Question 2

SEM analyses were conducted to examine the extent to which fidelity of implementation (use of *RC* practices) mediated the relation between assignment to the *RC* approach condition and achievement outcomes. The same child- and school-level covariates described in Table 4 were included in models described in Table 5. Fidelity was positively related to fifth-grade math and reading (β = .26, $p < .01$; β = .30, $p < .01$, respectively). Treatment related positively to fidelity (β = 1.73, $p < .01$). School-level percentage FRPL related positively, and teachers' years of experience related negatively to fidelity.

The direct effect of random assignment to treatment was negative for both outcomes (β = –.60, $p < .01$ for math; β = –.58, $p < .01$ for reading), indicating that, controlling for fidelity, students in *RC* schools showed lower math and reading achievement than those in the control schools. Overall tests of the indirect effects of treatment through fidelity on fifth-grade math and reading achievement were positive and significant (β = .44, $p <$ .01; β = .52, $p < .01$, respectively), indicating that random assignment to

Table 5
Mediation Model of Treatment on Fifth-Grade Math and Reading Achievement Through Fidelity

	Math Achievement			Reading Achievement		
Variable Name	*B*	*SE*	β	*B*	*SE*	β
Child level						
Gender (female)	11.05*	4.40	0.16	13.44**	3.00	0.21
Free/reduced price lunch (FRPL) status	–11.54**	4.40	–0.16	–12.47**	3.75	–0.19
English language learner status	–5.15*	2.21	–0.10	–7.63**	1.57	–0.17
Plain English test	–39.28**	11.75	–0.55	–28.46**	9.82	–0.45
Initial math achievement	1.09**	0.11	0.64	0.74**	0.05	0.48
School level						
Treatment status	–42.68**	10.64	–0.60	–37.25**	6.97	–0.58
Fidelity	19.08**	6.70	0.26	19.25**	3.53	0.30
Years of experience	1.32	0.84	0.07	0.32	0.51	0.02
Percentage of master's degrees	–7.32	19.71	–0.01	–0.19	13.86	–0.00
Percentage FRPL (school level)	–0.38	0.20	–0.12	–0.40**	0.12	–0.15
Initial math achievement (school level)	–0.33	0.30	–0.07	0.63**	0.18	0.16
Predictors of fidelity						
Treatment status	1.71**	0.17	1.73	1.71**	0.17	1.73
Years of experience	–0.04**	0.01	–0.14	–0.04**	0.01	–0.14
Percentage of master's degrees	–0.26	0.36	–0.04	–0.26	0.36	–0.04
Percentage FRPL (school level)	0.01*	0.00	0.24	0.01*	0.00	0.24
Initial math achievement (school level)	0.00	0.01	0.00	0.00	0.01	0.00
Treatment » Fidelity » Math	32.53**	11.35	0.44	32.81**	8.49	0.52

Note. R^2 = .88 for fidelity of implementation. R^2 = 0.35 for math achievement. R^2 = 0.45 for reading achievement. Child-level initial math achievement is group-mean centered.
*p < .05. **p < .01.

the *RC* condition caused increased fidelity, which in turn was associated with improvements in both math and reading achievement. Examination of effect size gains (based on standardized coefficients) suggests that being in an *RC* school with one *SD* higher levels of fidelity related to a .44 *SD* gain in math scores and a .52 *SD* gain in reading scores. The presence of the positive indirect effect needs to be interpreted in the context of the negative direct effect (β = –.60 for math; –.58 for reading, $p < .01$). Therefore, the total effects of random assignment (indirect effects of treatment plus the direct effect of treatments) are small and nonsignificant for both outcomes.

Partial residual plots (Larsen & McCleary, 1972) were created to display results graphically. Figure 3 shows partial residual plots for fifth-grade achievement and fidelity. The figure shows that all intervention schools

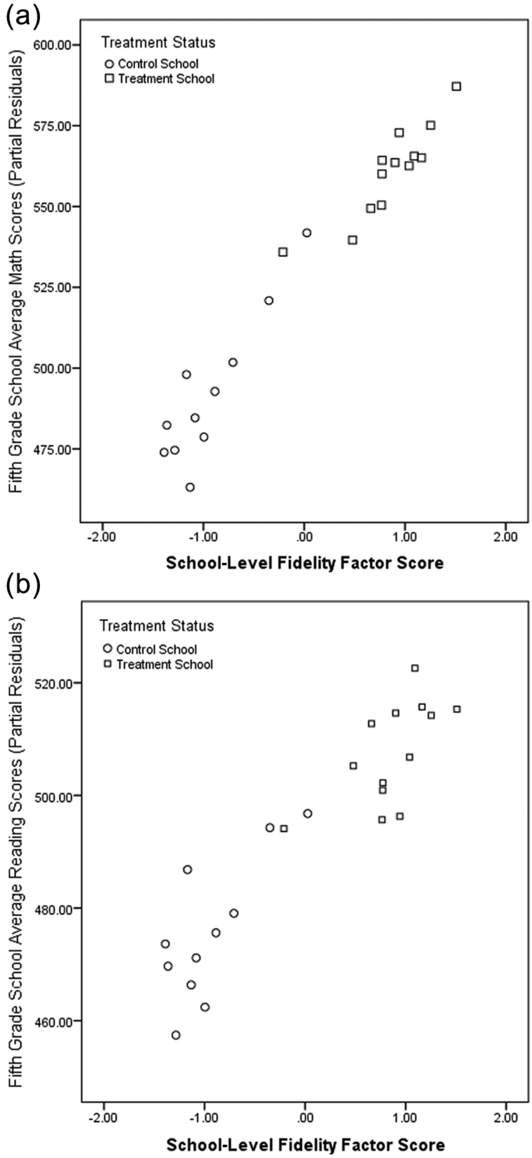

Figure 3. **Partial residual plot of school average fifth-grade math (a) and reading (b) scores versus fidelity.**

Note. All school level covariates are accounted for in the fifth-grade residuals.

except one showed higher fidelity and fifth-grade math and reading achievement compared to control schools, after controlling for covariates.

Research Question 3

Analyses were conducted to test the possibility of moderated mediation, such that the relation between treatment, fidelity, and the outcomes differed depending on students' eligibility for FRPL and students' initial math achievement (<25th percentile). Separate SEM models were conducted for each potential moderator for math and reading. Models allowed the paths from treatment status to fidelity, treatment status to achievement, and fidelity to achievement to vary depending on the moderators. Results showed no significant moderation effects for FRPL for either outcome. Initial achievement showed significant moderated mediation in relation to math achievement. The relation between fidelity and math achievement was stronger for students with low initial achievement. As a result, the overall mediation (*RC* to fidelity to math achievement) was significant for low-achieving students but not higher achieving students. Examination of effect size gains (based on standardized coefficients) suggests that being in an *RC* school with one *SD* higher levels of *RC* fidelity related to a .89 *SD* gain in math scores for students low in initial achievement but only a .25 *SD* gain in math scores for students higher in initial achievement. Initial achievement did not moderate any of the paths in the model predicting reading achievement. See Table 6.

Discussion

We discuss findings from a clustered randomized controlled trial on the *RC* approach and achievement in relation to three foci: (a) the impact of the *RC* approach over 3 years on students' reading and math achievement, (b) the extent to which fidelity of implementation mediated the relations of treatment assignment (intervention vs. control) with reading and math achievement, and (c) the extent to which paths in the mediation models were moderated by FRPL status and low initial math achievement.

Impact of the *RC* Approach

Random assignment to intervention versus control schools did not impact student achievement outcomes. Although the result was counter to our hypothesis, other RCT results linking SEL interventions to SEL outcomes are comparably lackluster (SACD, 2010). Several explanations are plausible. Treatment was randomized at the school level, and thus, analyses were also conducted at the school level. The a priori power analysis suggested sufficient statistical power to detect small to medium effects (.85 assuming ICC values of .10); however, this number of schools was far from ideal. The

Table 6

Moderated Mediation Model of Treatment on Fifth-Grade Math and Reading Achievement Through Fidelity for Low (<25%-ile) versus Higher (>25%-ile) Initial Achievement

	Math Achievement			Reading Achievement		
Variable Name	*B*	*SE*	β	*B*	*SE*	β
Child level						
Gender (female)	12.05**	4.08	0.17	13.71**	3.05	0.21
Free/reduced price lunch (FRPL) status	–10.92*	4.83	–0.15	–12.06**	4.00	–0.19
English language learner status	–3.06	2.05	–0.06	–4.92	4.64	–0.11
Plain English test	–32.64**	11.64	–0.46	–26.12*	10.12	–0.41
Initial math achievement (group-mean centered)	0.80**	0.12	0.46	0.64**	0.04	0.41
In lowest 25th percentile of Stanford 10	–25.61	19.07	–0.36	–15.16	10.14	–0.24
School level						
Treatment status	–25.78†	14.26	–0.36	–34.39**	6.97	–0.54
Fidelity	10.46	7.94	0.14	17.92**	3.63	0.28
Years of experience	1.33	0.84	0.07	0.27	0.53	0.02
Percentage of master's degrees	–5.12	20.55	–0.01	0.49	14.69	0.00
Percentage FRPL (school level)	–0.38*	0.19	–0.12	–0.41**	0.12	–0.15
Initial math achievement (school level)	–0.59†	19.07	–0.13	–15.16	10.14	–3.77
Cross-level interactions						
Low Achieving × Treatment Status	–42.61	30.61	–0.60	–3.31	18.16	–0.05
Low Achieving × Fidelity	26.42†	14.35	0.37	2.80	7.79	0.04
Predictors of fidelity						
Treatment status	1.74**	0.16	1.76	1.74**	0.16	1.76
Years of experience	–0.03**	0.01	–0.11	–0.03**	0.01	–0.11

(continued)

Table 6 **(continued)**

	Math Achievement			Reading Achievement		
Variable Name	*B*	*SE*	β	*B*	*SE*	β
Percentage of master's degrees	–0.25	0.28	0.02	–0.25	0.28	0.02
Percentage FRPL (school level)	0.01*	0.00	0.24	0.01*	0.00	0.24
Initial math achievement (school level)	0.00	0.05	0.00	0.00	0.01	0.00
In lowest 25th percentile of Stanford 10	0.00	0.05	0.00	–0.01	0.07	0.01
Low Achieving × Treatment Status	–0.04	0.07	–0.04	–0.03	0.07	–0.03
Conditional indirect effects						
Treatment » Fidelity » Outcome (Low Achieving, ≤ 25th percentile)	63.00**	22.91	0.89	33.44*	14.59	0.52
Treatment » Fidelity » Outcome (Higher Achieving >25th percentile)	18.23	13.08	0.25	31.23**	8.11	0.49
Total effects						
Total effect (low achieving)	–5.39	13.60	–0.08	–2.26	7.97	–0.04
Total effect (high achieving)	–7.55	6.23	–0.10	–3.14	5.54	–0.05

Note. R^2 = 0.91 for fidelity. R^2 = 0.29 for math achievement. R^2 = 0.46 for reading achievement.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

real effect size could be smaller than anticipated and the study conducted here may have lacked sufficient statistical power to detect effects. Second, other constructs believed to be integral to our logic model (e.g., student motivation and engagement) may not adequately translate into outcomes measured by state standardized achievement tests (Duckworth, Quinn, & Tsukayama, 2012). A third reason pertains to fidelity of implementation of the *RC* approach. We describe this point carefully because the explanation is consistent with accumulated research pointing to the importance of high-quality implementation of SEL interventions for predicting child outcomes (Domitrovich et al., 2008) and also because this point has the greatest relevance for practitioners making decisions about use of the *RC* approach.

Role of Fidelity of Implementation

Training in the *RC* approach produced large changes in teacher practices. Measures of both frequency and adherence to the *RC* approach indicate the presence of treatment differentiation between intervention and control schools (effect size *d* values: 2.47–3.88). We used teacher-report and observational fidelity measures to determine whether the intervention was delivered and to assess the quality of that delivery (O'Donnell, 2008). Fidelity results were consistent with the program logic model, permitting us to investigate fidelity as a mechanism for explaining potential outcomes. An important nuance emerged from fidelity analyses. Results revealed evidence of high but variable intervention uptake among intervention schools and low but variable *RC* practices in control schools. It is not surprising that *RC* practices were present in control schools given the common origins between *RC* practices and "best practices" in teaching and intervention diffusion (Hulleman & Cordray, 2009).

Receiving the *RC* training produced gains in *RC* practices, which in turn was associated with improvements in both math and reading achievement. The effect size (based on standardized coefficients) between use of *RC* practices and student achievement was .26 and .30 for math and reading, respectively. Findings from the Durlak et al. (2011) meta-analysis report effect sizes (Hedge's *g*) of .27 for SEL interventions on achievement and describe lower effect sizes in the presence of fidelity problems. Results from a meta-analysis of 76 meta-analytic papers on educational interventions give benchmark effects for the elementary grades ranging from .22 to .23 (Hill, Bloom, Black, & Lipsey, 2008). The effect size values for the association from *RC* practices to achievement are comparable to those for SEL and other educational interventions for this age group.

The fidelity of implementation findings raises an enigma for practitioners and policymakers making decisions about the *RC* approach. On one hand, *RC* practices link positively to achievement. Randomization to the *RC* condition causes use of *RC* practices, which in turn relates to

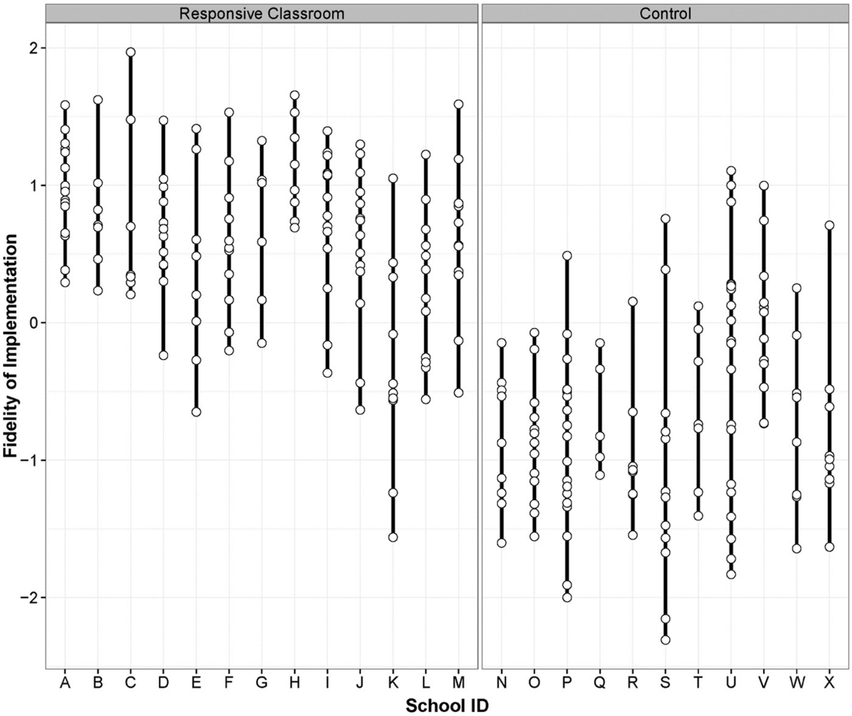

Figure 4. **Fidelity of implementation at intervention and comparison schools.**

increased student achievement. On the other hand, the direct effect of *RC* training to achievement is negative in models that include fidelity of implementation. Hypothetically, what this means is that given two schools (one intervention and one control) with equivalent levels of *RC* practices, students in the control condition will outperform the intervention condition. Yet realistically, there are only 2 of 24 schools in the present sample that fits with this hypothetical example. In the remaining 22 cases, intervention schools use more *RC* practices than control schools, producing the exact conditions associated with student achievement gains.

The enigma here stems from the common problem of treatment heterogeneity in cluster randomized controlled trials (Cook, 2005). Schools vary in their use of *RC* practices, as shown in Figure 4. Such variability calls attention to factors that predict successful uptake of *RC* practices. As one example, mixed-methods research from RCES shows the importance of school leadership. Teachers' perception of their principals' "buy-in" to the intervention related to teachers' use of *RC* practices. Teachers questioned using the *RC* practices if they thought their principals adopted the approach to acquire funds or prestige. Also, teachers used fewer practices if their principals adopted programs that conflicted with *RC* principles but used more practices if their principal made structural accommodations to engender their use (Wanless, Patton, Rimm-Kaufman, & Deutsch, 2013). Other SEL work (e.g., PATHS) shows the importance of administrative support in predicting higher quality of implementation of the intervention (Ransford, Greenberg, Domitrovich, Small, & Jacobson, 2009). Schools are social systems in which leadership decisions influence teachers, students, and psychological experiences within the school (Bryk, Sebring, Allensworth, Easton, & Luppescu, 2010). Administrators who convey clear support for the *RC* approach and indicate commitment to long-term adoption may enhance teachers' use of *RC* practices.

Other factors are likely to be important predictors of uptake of the *RC* approach. Teachers' psychological state (e.g., burnout, self-efficacy) and their perceived value of the program have been implicated in schoolwide SEL adoption (Brackett, Reyes, Rivers, Elbertson, & Salovey, 2011; Ransford et al., 2009). Effective coaching supports implementation of SEL practices (Ransford et al., 2009). Teachers' and administrators' comfort, commitment, and culture to SEL (Brackett et al., 2011) and organizational health of the school community (Bradshaw, Koth, Thornton, & Leaf, 2009) related to higher implementation of practices or moderate the link between teachers' psychological characteristics and use of practices. Teachers' advice networks (i.e., access to intervention expertise, conversation with other teachers about practices, pressure from other teachers, common sources of mentorship among teachers) contribute to the adoption of new practices (Frank, Zhao, & Borman, 2004; Neal, Neal, Atkins, Henry, & Frazier, 2011).

Assignment to the intervention condition without use of *RC* practices relates lower student achievement, a finding requiring explanation. One possibility is that learning the *RC* practices disrupted typical high-quality instruction used before training in the *RC* approach. Critics of the *RC* approach suggest that time spent engaged in *RC* practices (e.g., Morning Meeting, which can take 20 minutes per day) detracts from instructional time. In contrast, developers of the *RC* approach suggest that time allocated to *RC* practices has benefits later in the day and year (Denton & Kriete, 2000). Other analyses on this data set examined the *RC* approach and quality and quantity of math instruction. *RC* and comparison schools did not differ in the amount of time spent in math instruction on three typical days during the year and third-grade math teachers at *RC* schools showed improved, not diminished, inquiry-based math instruction (Ottmar, Rimm-Kaufman, Berry, & Larsen, 2013). Although limited to math only, existing results on a subset of teachers in RCES do not show reduced time or quality.

Another possibility is that there are attributes of the schools that predict both decreased uptake of the intervention *and* lower student achievement. For example, if school leaders lack leadership skills, teacher may have difficulty creating priorities in the face of competing demands on their time. In turn, poor leadership may lead to both low intervention uptake and low achievement. It is not possible to determine the direction of causality—poor leadership may produce low *RC* implementation and low achievement or it may be harder to implement *RC* practices and lead effectively in conditions with very low student achievement.

Other explanations are also plausible. Within a single school, some teachers may show more uptake of the intervention than others. The positive indirect path would be due to the teachers who uptake the intervention more and the negative direct effects would be due to teachers who do not use *RC* practices. Yet another explanation is that the *RC* approach has both positive and negative components. The positive components may be well represented in the fidelity measure while the negative components were not. Further research using mixed-methods approaches will help explain these findings.

Subgroup Analysis

Findings showed that the relations among treatment assignment, fidelity of implementation, and reading and math achievement were equivalent regardless of whether students qualified for FRPL or not. This finding is evident after controlling for the poverty composition of the school. Replication of the work is needed to understand the generalizability of findings to other schools in different settings (e.g., extremely high poverty schools).

Enrollment at an *RC* school with high implementation conveyed greater advantages to students' math achievement if students were initially in the bottom math achievement quartile. The differential contribution of *RC* practices draws attention to two possible mechanisms—student skill development and teacher capacity. In relation to student skills, the *RC* approach is designed to improve students' self-regulatory and attentional skills, skills that may undergird student math achievement (Duncan et al., 2007; Li-Grining et al., 2010). Pertaining to teachers' capacity, the *RC* approach is designed to improve a teacher's ability to understand and meet students' *individual* needs. One *RC* principle states "Knowing the children we teach—individually, culturally, and developmentally—is as important as knowing the content we teach" (NEFC, 2007, p. 3). Various *RC* practices such as Teacher Language and Academic Choice may improve teachers' capacity to work with low-achieving students.

Limitations

Several limitations require mention. First, the quality of results may be limited by the selected outcome measure. The state standardized achievement test has been criticized because of the presence of ceiling effects and because reliability is higher near policy-relevant cut-points (e.g., those defining proficiency) compared to other points in the scale (May, Perez-Johnson, Haimson, Sattar, & Gleason, 2009). Further, the math standardized achievement tests, although critical to policy decisions, tap procedural understanding and test-taking ability, not conceptual understanding of mathematics. Statistical approaches, such as censored data techniques, were used to compensate for some test limitations. Second, timing and resource constraints influenced data collection decisions. Because of timing constraints, third-grade teachers were observed in their first year of implementation (after *RC* 1) whereas fourth- and fifth-grade teachers were observed in their second year of implementation (after *RC* 2). Although not ideal, NEFC states that many teachers trained in the *RC* approach only attend *RC* 1. Because of project priorities, math scores were used as indicators of low initial achievement and 60% of the observational fidelity measures were gathered during math instruction; thus, findings may be more precise for math than reading effects. This limitation must be considered in interpreting the *RC* practices to achievement link for students with low initial math achievement. However, this concern is partially ameliorated in that 40% of the observational measures and all of the teacher-report measures of fidelity of implementation were not gathered during math instruction. Third, adopting interventions such as the *RC* approach involves a long process of teacher change ranging from 3 to 5 years. Data collected in the present study were gathered during teachers' first and second years of *RC* implementation, early in the process of *RC* adoption.

Implications and Future Directions

This study is among the few randomized controlled trials examining the effect of SEL interventions on student achievement. Findings indicate that student gains hinge on actual use of recommended classroom practices. This study considered a question school administrators and teachers ask: Will we diminish children's academic achievement if we place increased emphasis on (and allocate more time toward) children's social and emotional learning? Findings show that using the *RC* approach did not diminish achievement, and if the *RC* practices were used as intended, students showed achievement gains. Further, use of *RC* practices was linked to enhanced achievement in children who were initially low in math achievement, a point in need of replication given that the findings do not permit causal inference.

Results direct attention toward next steps. Treatment heterogeneity among the schools suggests the importance of examining teacher-level implementation in future work. Mixed-methods and qualitative work is needed. For instance, findings raise important questions about the supports and barriers that predict teacher and school implementation of *RC* practices as well as questions about the process of teacher change involved as teachers implement the *RC* approach. Yet another needed step is work that considers students' exposure to *RC* practices. For instance, do the contributions of *RC* practices become more pronounced after 2 or 3 years of student exposure to *RC* compared to after only a single year? Do some students find *RC* practices more engaging than others and thus participate more actively and benefit more substantially compared to other students? Training in the *RC* approach appears to be very effective in producing teacher change. Other interventions that prove less effective at creating teacher change may consider referring to the NEFC training model. Following students into middle and high school years and expanding outcomes to include social and emotional constructs may provide insight into potential enduring contributions of *RC* practices.

Notes

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305A070063, R305B060009, and R305B040049 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. We gratefully acknowledge the contributions of Robert Berry, Gretchen Bukowick, Claire Cameron, Nikki Fedoravicius, Tom Fruscello, Roxann Kriete, Abigail Moncrief, Kerry O'Grady, Erin Ottmar, Christine Patton, Jennifer Williams, and the administrators, research and evaluation staff, teachers, and students in our collaborating school district.

[1]A moderator variable is a variable (or condition) that changes the association between an independent variable and a dependent variable. A mediator variable is a third variable that explains the mechanism underlying the relation between an independent and a dependent variable. Moderated mediation occurs when a mediation effect differs depending on the level of the moderator variable (e.g., eligibility for free/reduced price lunch) (Fairchild & McQuillin, 2010).

References

Abbott, R. D., O'Donnell, J., Hawkins, J. D., Hill, K. G., Kosterman, R., & Catalano, R. F. (1998). Changing teaching practices to promote achievement and bonding to school. *American Journal of Orthopsychiatry*, *68*(4), 542–552.

Abry, T., Brewer, A. J., Nathanson, L., Sawyer, B., & Rimm-Kaufman, S. E. (2010). *Classroom Practices Observation Measure*. Unpublished measure, University of Virginia, Charlottesville, VA.

Abry, T., Rimm-Kaufman, S. E., Larsen, R. A. A., & Brewer, A. J. (2013). The influence of fidelity of implementation on teacher-student interaction quality in the context of a randomized controlled trial of the *Responsive Classroom* approach. *Journal of School Psychology*, *51*, 437–453. doi:http://dx.doi.org/10/1016/j.jsp.2013.03.001

Allison, P. D. (2002). Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology*, *55*(1), 193–196.

Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal*, *103*(1), 51–73.

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*(3), 243–262.

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, *32*(3), 627–658.

Bickman, L., Riemer, M., Brown, J. L., Jones, S. M., Flay, B. R., Li, K. K., & . . . Massetti, G. (2009). Approaches to measuring implementation fidelity in school-based program evaluations. *Journal of Research in Character Education*, *7*(2), 75–101.

Bloom, H. S., & Michalopoulos, C. (2010). When is the story in the subgroups? *Prevention Science*, 1–10.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2011). Assessing teachers' beliefs about social and emotional learning. *Journal of Psychoeducational Assessment*. Advance online publication. doi:10.1177/0734282911424879

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2010). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*, 218–224.

Bradshaw, C. P., Koth, C. W., Thornton, L. A., & Leaf, P. J. (2009). Altering school climate through school-wide positive behavioral interventions and supports: Findings from a group-randomized effectiveness trial. *Prevention Science*, *10*(2), 100–115.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, *12*(3), 133–148.

Bradshaw, C. P., Zmuda, J. H., Kellam, S. G., & Ialongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology*, *101*(4), 926–937. doi:10.1037/a0016586

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Century, J. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *The American Journal of Evaluation*, *31*(2), 199–218. doi:10.1177/1098214010366173

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*(4), 415–427.

Collaborative for Academic Social and Emotional Learning (CASEL). (2013). *CASEL guide: Effective social and emotional learning programs*. Retrieved from www.casel.org/guide.

Conduct Problems Prevention Research Group. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, *78*(2), 156–168.

Cook, T. D. (2005). Emergent principles for the design, implementation, and analysis of cluster-based experiments in social science. *The Annals of the American Academy of Political and Social Science*, *599*(1), 176–198.

Cordray, D. S., & Pion, G. M. (Eds.). (2006). *Treatment strength and integrity: Models and methods*. Washington, DC: American Psychological Association.

Crosnoe, R., & Cooper, C. E. (2010). Economically disadvantaged children's transitions into elementary school: Linking family processes, school contexts, and educational policy. *American Educational Research Journal*, *47*(2), 258–291.

Crosnoe, R., Morrison, F., Burchinal, M., Pianta, R., Keating, D., Friedman, S. L., & Clarke-Stewart, K. A. (2010). Instruction, teacher-student relations, and math achievement trajectories in elementary school. *Journal of Educational Psychology*, *102*(2), 407–417. doi:10.1037/a0017762

Denton, P. (2005). *Learning through academic choice*. Turner Falls, MA: Northeast Foundation for Children.

Denton, P., & Kriete, R. (2000). *The first six weeks of school*. Turner Falls, MA: Northeast Foundation for Children.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., & . . . Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, *1*(3), 6–28.

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*(2), 439–451.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & . . . Brooks-Gunn, J. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446.

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 on social and emotional learning. *Child Development*, *82*(1), 405–432. doi:10.1111/j.1467-8624.2010.01564.x

Dusenbury, L., Zadrazil, J., Mart, A., & Weissberg, R. (2011). *State learning standards to advance social and emotional learning: The state scan of social and emotional learning standards, preschool through high school*. Retrieved from http://casel.org/wp-content/uploads/Forum-Brief-on-the-State-Scan-5-10-2011.pdf.

Elliott, S. N. (1999). A *multi-year evaluation of the* Responsive Classroom *approach: Its effectiveness and acceptability in promoting social and academic competence*. Retrieved from http://www.responsiveclassroom.org/PDF_files/final_report.pdf.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*(3), 430–457.

Evans, R. (2001). *The human side of school change*. 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*(1), 53–84.

Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: The case of computer technology in schools. *Sociology of Education*, *77*, 148–171.

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*(1), 70–95.

Greenberg, M. T. (2010). School-based prevention: current status and future challenges. *Effective Education*, *2*(1), 27–52.

Harcourt Educational Measurement (2003). *Stanford Achievement Test Series, Tenth Edition, Primary 2*. New York: Harcourt Inc.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177.

Hughes, J., & Kwok, O. M. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology*, *99*(1), 39–51. doi:10.1037/0022-0663.99.1.39

Hulleman, C., & 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*(1), 88–110.

Jones, S. M., Brown, J. L., & Aber, J. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An Experiment in translational developmental research. *Child Development*, *82*, 533–554.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, *100*(4), 851–868.

Klein, J., & Moeschberger, M. (1997). *Survival analysis: Techniques for censored and truncated regression*. New York, NY: Springer-Verlag.

Knowlton, L. W., & Phillips, C. C. (2008). *The logic model guidebook: Better strategies for great results*. Thousand Oaks, CA: Sage.

Kriete, R. (2002). *The Morning Meeting book*. Turner Falls, MA: Northeast Foundation for Children.

Larsen, W. A., & McCleary, S. J. (1972). The use of partial residual plots in regression analysis. *Technometrics*, *14*(3), 781–790.

Lee, V., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal*, *36*, 907–945.

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, *46*(5), 1062–1077. doi:10.1037/a0020066

Little, R. J. A., & Rubin, D. (1987). *Statistical analysis with incomplete data*. New York, NY: John Wiley & Sons.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables* (*Vol. 7*). Thousand Oaks, CA: Sage.

May, H., Perez-Johnson, I., Haimson, J., Sattar, S., & Gleason, P. (2009). *Using state tests in education experiments: A discussion of the issues* (NCEE 2009-013). Washington, DC: National Center for Education Evaluation and Regional Assistance.

Murray, D. M., Hannan, P. J., & Baker, W. L. (1996). A Monte Carlo study of alternative responses to intraclass correlation in community trials: Is it ever possible to avoid Cornfield's penalties? *Evaluation Review*, *20*(3), 313–337.

Muthén, L., & Muthén, B. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén and Muthén.

Nathanson, L., Sawyer, B., & Rimm-Kaufman, S.E. (2007a). *Classroom Practices Teacher Survey*. Unpublished instrument, University of Virginia, Charlottesville, VA.

Nathanson, L., Sawyer, B., & Rimm-Kaufman, S.E. (2007b). *Classroom Practices Frequency Survey*. Unpublished instrument, University of Virginia, Charlottesville, VA.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Neal, J. W., Neal, Z. P., Atkins, M. S., Henry, D. B., & Frazier, S. L. (2011). Channels of change: Contrasting network mechanisms in the use of interventions. *American Journal of Community Psychology*, *47*, 277–286.

Northeast Foundation for Children. (2007). *Responsive classroom: Level 1 resource book*. Turners Falls, MA: Northeast Foundation for Children.

Northeast Foundation for Children. (2009). *Responsive classroom: Level 2 resource book*. Turners Falls, MA: Northeast Foundation for Children.

Northeast Foundation for Children (2014a). Good teaching changes the future. Retrieved February 12, 2014 from: http://www.responsiveclassroom.org/sites/default/files/pdf_files/RC_approach_White_paper.pdf

Northeast Foundation for Children (2014b). Principles and practices of responsive classroom. Retrieved February 12, 2014 from: http://www.responsiveclassroom.org/principles-and-practices-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*(1), 33–84.

Ottmar, E. R., Rimm-Kaufman, S. E, Berry, R. Q., & Larsen, R. A. A. (2013). Results from a randomized controlled trial: Does the *Responsive Classroom* approach impact the use of standards-based mathematics teaching practices? *Elementary School Journal*, *113*(3), 434–457. doi:10.1086/668810

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, *45*(2), 365–397.

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*(4), 510–532.

Raver, C. C. (2012). Low-income children's self-regulation in the classroom: Scientific inquiry for social change. *American Psychologist*, *67*(8), 681–689. doi:10.1037/a0030085

Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *School Psychology Review*, *41*(1), 82–99.

Rimm-Kaufman, S. E., & Chiu, Y. I. (2007). Promoting social and academic competence in the classroom: An intervention study examining the contribution of the *Responsive Classroom* approach. *Psychology in the Schools*, *44*(4), 397–413.

Rimm-Kaufman, S. E., Fan, X., Chiu, Y. I., & 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.

Rimm-Kaufman, S. E., Storm, M. D., Sawyer, B. E., Pianta, R. C., & La Paro, 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.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*(2), 268–302.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Shouse, R. C. (1996). Academic press and sense of community: Conflict, congruence, and implications for student achievement. *Social Psychology of Education*, *1*(1), 47–68.

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., & Li, K. K. (2009). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster-randomized, controlled trial. *Journal of Research on Educational Effectiveness*, *3*, 26–55.

Social and Character Development Research Consortium. (2010). *Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children* (NCER 2011–2001). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Virginia Department of Education. (2008). *Virginia standards of learning technical report: 2008-2009 administration cycle*. Retrieved from http://www.doe.virginia.gov/testing/test_administration/technical_reports/sol_technical_report_2008-09_administration_cycle.pdf.

Virginia Department of Education. (2010). *Virginia standards of learning assessment: Test blueprint, grade 4 mathematics*. Retrieved from http://www.doe.virginia.gov/testing/sol/blueprints/mathematics_blueprints/2009/blueprint_math4%20.pdf.

Votruba-Drzal, E. (2006). Economic disparities in middle childhood development: Does income matter? *Developmental Psychology*, *42*(6), 1154–1167. doi:10.1037/0012-1649.42.6.1154

Wanless, S. B., Patton, C. S., Rimm-Kaufman, S. E., & Deutsch, N. L. (2013). Setting-level influences on implementation of the *Responsive Classroom* approach. *Prevention Science*, *14*(1), 40–51. doi:10.1007/s11121-012-0294-1

Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention. *American Psychologist*, *67*(4), 272–284. doi: 10.1037/a0028015

Manuscript received January 2, 2013
Final revision received September 5, 2013
Accepted November 2, 2013