



Published in final edited form as:

Early Educ Dev. 2013 ; 24(8): . doi:10.1080/10409289.2013.755457.

Sustaining High Quality Teaching and Evidence-based Curricula: Follow-up Assessment of Teachers in the REDI Project

Karen L Bierman, PhD,

Department of Psychology, The Pennsylvania State University

Rebecca M. Sanford DeRousie, Ph.D., Brenda Heinrichs, M.S., Celene E. Domitrovich, Ph.D., Mark T. Greenberg, Ph.D., and Sukhdeep Gill, Ph.D.

Abstract

Recent research has validated the power of evidence-based preschool interventions to improve teaching quality and promote child school readiness when implemented in the context of research trials. However, very rarely are follow-up assessments conducted with teachers in order to evaluate the maintenance of improved teaching quality or sustained use of evidence-based curriculum components after the intervention trial. In the current study, we collected follow-up assessments of teachers one year after their involvement in the REDI research trial to evaluate the extent to which intervention teachers continued to implement the REDI curriculum components with high-quality, and to explore possible pre-intervention predictors of sustained implementation. In addition, we conducted classroom observations to determine whether general improvements in the teaching quality of intervention teachers (relative to control group teachers) were sustained. Results indicated sustained high-quality implementation of some curriculum components (the PATHS curriculum), but decreased implementation of other components (the language-literacy components). Sustained intervention effects were evident on most aspects of general teaching quality targeted by the intervention. Implications for practice and policy are discussed.

Early education has become a primary strategy for reducing the disparities in school readiness associated with poverty, with the goal of fostering the long term academic success of economically-disadvantaged children (Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002; Ramey & Ramey, 2004). During the past 10–15 years, a number of rigorous randomized trials have demonstrated that enhancing preschool curricula with evidence-based practices improves school readiness in areas of emergent literacy and numeracy skills (Lonigan, 2006; Sarama & Clements, 2009), and in areas of social-emotional development and learning behaviors (Bierman, Domitrovich, & Darling, 2009; Joseph & Strain, 2003). In addition, professional development activities (e.g., workshops, coaching, videotaped feedback) have proven effective at improving the quality of preschool teaching and teacher-student interactions, which in turn improve child school readiness (Fox & Hemmeter, 2009; Pianta, Mashburn, Downer, Hamre, & Justice, 2008).

Although a number of evidence-based interventions have demonstrated the potential to improve teaching quality and enhance children's school readiness, very few preschool studies have collected follow-up assessments on teachers after the intervention trial, in order to assess the maintenance of gains in teaching quality and evaluate the sustained implementation of the curriculum. These are important issues, given that evidence-based interventions can have a significant impact on child outcomes over time only if they are maintained with high-quality implementation (Durlak & Dupree, 2008). Research in other areas of prevention and educational interventions with older children suggest that the goal of sustained, high-fidelity program implementation is difficult to attain; evidence-based

interventions often fade quickly or are replaced by alternative programs after initial funding and implementation support is withdrawn (Elias, 2004; Florian, 2001).

The present study contributes to this emerging research literature in early education interventions by examining follow-up assessments of teachers who participated in the randomized trial of the Head Start REDI (Research-based, Developmentally Informed) project, which included evidence-based preschool curriculum components and professional development support to enhance teaching quality. One year after the research trial, we revisited teachers who had participated in the research trial in the intervention or “usual practice” control group and examined the quality of their teaching in areas of REDI focus. Within the intervention group, we also evaluated the sustained high-quality implementation of two types of intervention components (social-emotional and literacy programming). Teacher and program characteristics measured prior to the research trial, along with the quality of program implementation during the research trial, were evaluated as predictors of sustained REDI program implementation at the follow-up assessment.

The Importance of Collecting Follow-Up Assessments of Teaching Practices

Typically, evidence-based interventions are introduced into school settings during an initial, time-limited period when funding is allocated to support professional development and implementation fidelity. During this initial phase, teachers receive program materials and workshop training, and often receive on-going coaching as they implement the intervention. Child outcomes are collected at the end of this intervention trial. Preschool intervention research often includes follow-up assessments of the children involved, to determine the longevity of intervention benefits. However, follow-up assessments of teachers are rare in the early education research literature (see also Florian, 2001). Schools expect that teachers will maintain the skills gained in professional development and that curricular enhancements will become institutionalized and integrated into their permanent practice, particularly if they are valued by the teachers and program – yet rarely are levels of sustained impact actually measured.

Studies examining the sustainability of evidence-based preschool interventions are extremely rare; however, several studies have examined the sustainability of school-based intervention at the elementary or secondary school level, providing a basis for conceptualizing the challenge. As reviewed briefly here, these studies on older school populations suggest that sustained improvements are difficult to attain, as most evidence-based interventions decline in quality over time (Elliott & Mihalic, 2004; Han & Weiss, 2005). For example, Elias (2004) examined the sustainability of evidence-based social-emotional curricula in schools, and found that only 6 of 14 sites were still using the programs 5 years after their introduction. Even when programs continue to use evidence-based curricula, changes often occur in the implementation quality, reducing the positive impact on child outcomes (Dane & Schneider, 1998; Durlak & Dupree, 2008). For example, a follow-up study of Life Skills Training, a school-based prevention program, found that teachers rarely implemented the program as written (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005). Whereas adaptations may sometimes enrich the program and enhance impact, often adaptations involve reductions in intervention intensity or modifications of core program components, thereby limiting impact (Elliott & Mihalic, 2004).

In an educational intervention, program impact is affected not only by the adherence to the intervention components (e.g., teaching the program content) but also by the quality with which intervention lessons are delivered (Dane & Schneider, 1998). The present follow-up study used coach ratings to assess the on-going implementation of the REDI curriculum

components in the intervention classrooms. Coaches made these ratings after reviewing teacher lesson plans, talking with teachers about their use of the program (frequency of use; degree of modification), and after they observed lessons being taught. As such, the measure of program sustainability used here was a sum score that reflected teacher-reported intervention adherence, as well as coach observations of implementation quality.

In addition to the delivery of the specific curriculum lessons, the REDI intervention also utilized professional development activities designed to improve the quality of more general teaching practices, particularly teacher's emotional support, positive behavioral management strategies, and language use in the classroom. Attempts to improve these aspects of general teaching quality were motivated by early childhood research suggesting that these aspects of teaching quality play a primary role in fostering child skill development and school readiness (Burchinal et al., 2002; Pianta et al., 2008; Mashburn & Pianta, 2006). Whereas assessments of the sustained use of the intervention program could be measured only in the intervention classrooms, the more general aspects of teaching quality were observed in both the intervention and control classrooms. In addition, factors that might predict sustained use of the REDI curriculum components were explored.

Factors Influencing Sustained Programming

Previous research suggests that several factors may facilitate (or undermine) the sustainability of evidence-based programs. Conceptually, these factors can be grouped into two major categories. As described in more detail below, the first group are factors that reflect a program's *readiness for new program implementation*, which include: 1) the organizational climate, quality of supervisory support, and general morale of the teaching staff (Gottfredson, Jones, & Gore, 2002), 2) teacher knowledge and understanding of principles of early childhood education, reflected in their formal education levels (Vaughn, Klingner & Hughes, 2000), and 3) teacher skills, reflected in the pre-intervention quality of teaching practices (Gottfredson et al., 2002). The second group of factors has to do with the quality of teacher *experiences with program implementation* during its inaugural year, including teacher evaluations of the intervention (Han & Weiss, 2005), and the quality with which teachers initially learned and implemented the intervention (Han & Weiss, 2005; Vaughn et al., 2000).

Readiness for program implementation

Findings from comprehensive school reform studies suggest that disordered school and classroom environments impede the implementation of new reforms and thereby reduce positive program outcomes (Gottfredson et al., 2002). Limited organization capacity, in the form of poor staff morale, history of failed programs, and sense of resignation, has been associated with difficulty in implementing and sustaining innovations. In contrast, teachers may invest more effort in the sustained implementation of a new program when they feel supported and empowered by the program to do so (Greenberg, Domitrovich, Graczyk, & Zins, 2005; Scheirer, 2005; Dusenbury et al, 2003).

Teacher education and skill levels may also affect their readiness for sustained program implementation. In Head Start and other early childhood settings, teachers vary considerably in the amount of formal education they have, ranging from high-school education to bachelor or master's degrees. Teacher education has been linked with the overall quality of care in a number of early childhood studies (Blau, 2000; de Kruif, McWilliam, Ridley, & Wakeley, 2000), and may enable teachers to engage in more comprehensive and sophisticated implementation of research-based interventions and teaching practices (Whitehurst & Lonigan, 1998). In a prior study of the REDI program, Domitrovich and colleagues (Domitrovich, Gest, Jones, Gill, & Sanford DeRousie, 2010) found that teacher

education enhanced the quality of REDI program implementation during the research trial, suggesting that it might also facilitate sustained program use. In addition, a teacher who runs a well-managed and organized classroom may be better able to successfully navigate the challenges associated with implementing a new curriculum than a teacher who is struggling to maintain order or is less systematic in his/her approach to teaching (Hughes, Cavell, Meehan, Zhang & Collie, 2005). In addition, teachers who already have a solid foundation in core teaching skills, such as warm and sensitive responding, positive discipline practices, and rich, varied language use may find it easier to sustain the implementation of evidence-based curriculum components that utilize these skills.

Experiences with program implementation

In addition to their general skills, teacher's specific experiences with a particular intervention program may affect the degree to which and quality with which they sustain the use of that intervention. For example, teacher beliefs about the perceived effectiveness and acceptability of an intervention may affect their efforts at sustainability (Greenberg et al., 2005; Klinger, Vaughn, Hughes, & Arguelles, 1999; Scheier, 2005). When teachers perceive positive effects of a new teaching practice, they may be more willing to change their practices and embrace a new curriculum (Clarke & Hollingsworth, 2002; Han & Weiss, 2005). Curriculum acceptability (i.e., how user-friendly it is, the ease of fitting it into the day, and whether the materials are seen as age-appropriate and engaging) also appears important (Florian, 2001). Indeed, in the REDI project, teachers who perceived the intervention as congruent with their teaching style showed higher quality implementation during the research trial compared to those who did not perceive such congruence (Domitrovich et al., 2010).

In addition, the degree to which a teacher learns and effectively implements the intervention during the year of introduction and training support may affect the quality of sustained program use. Researchers have theorized that sustainability builds off of implementation, such that initial high-fidelity implementation improves the likelihood of program sustainability (Pluye, Potvin, & Menis, 2004; Scheirer, 2005). Moreover, the factors that predict and enhance initial implementation fidelity may also support sustained, high-fidelity implementation over time (Hans & Weiss, 2005). However, to date, no studies have examined directly the relationship between initial implementation quality of a preschool evidence-based program and the sustained use of the curriculum components or teaching practices.

The Current Study

A previously published study took an initial look at the sustainability of REDI program implementation during the year following the research trial (Sanford DeRousie & Bierman, 2012.) In that study, estimates of program sustainability were derived from three types of measures (e.g., teacher ratings, qualitative interviews with teachers, and REDI coach ratings). Results revealed variation in estimates of sustainability depending upon the method of measurement, as teachers reported higher levels of sustained program use than did coaches, probably because the coach ratings included judgments about the quality as well as quantity of sustained program use. Across the different measures and sources of input, higher levels of sustained use were documented for the social-emotional curriculum (the Preschool PATHS Curriculum) than for the language and emergent literacy curriculum components (interactive reading program, sound games, alphabet center).

This study extends the examination of the REDI follow-up assessments in two ways. First, this study explored factors that might serve as predictors of sustained high-quality curriculum use. Specifically, it explored the correlations between pre-intervention factors

reflecting readiness for program implementation along with intervention experiences during the inaugural year as predictors of sustained program implementation quality during the follow-up year. Given evidence from the prior study that the social-emotional and language-literacy intervention components were sustained at different levels, predictive correlations were examined separately for these two curriculum domains. It was anticipated that both readiness factors and experiences during the inaugural year of intervention would show associations with the sustained levels of implementation quality observed during the follow-up year.

This study also examined general teaching practices that were a focus of the REDI intervention, and compared the intervention and control teachers at the follow-up assessment to evaluate sustained intervention effects. Our hypothesis was that teachers would sustain the gains made in teaching quality during the intervention year, and because these gains reflected general strategies for interacting with children in the classroom, we anticipated that they would sustain across domains.

Method

Design Overview

Head Start classrooms in three Pennsylvania counties were stratified on county location, length of program (half or full-day), student demographics (proportion of minority students), and center size to assure even representation in the intervention and control conditions. Within each stratified group, classrooms were randomly assigned to condition. Pre-intervention data was collected in the spring prior to the implementation of the intervention, and follow-up data was collected two years later, a year after the research trial in which the intervention was introduced.

Participants

Participants included 37 lead classroom teachers who had both pre-intervention and follow-up assessment data. These teachers represented 82% of the original teachers in the REDI study; the other 7 teachers left Head Start prior to the collection of follow-up data. One teacher was male, the rest were female. Most teachers were Caucasian (85% of the intervention teachers; 82% of the control teachers). Each condition (intervention and control) also had two African American teachers and each had one Hispanic teacher. Most teachers had a 4-year college degree (I = 60%, C = 59%); others had a child development associate credential (I = 20%, C = 35%), or high school education or equivalent (I = 20%, C = 6%). Over half of the teachers had 11 or more years of teaching experience (I = 55%; C = 53%). One-quarter were new teachers, with 0–5 years of prior teaching experience (I = 25%, C = 23.5%), and the others had moderate levels of experience (6–10 years of experience; I = 20%, C = 23.5%).

Intervention Model

The intervention was delivered by classroom teachers and integrated into existing curricula (i.e., Creative Curriculum and High/Scope) used by the Head Start centers. To promote children's social-emotional development, REDI used the Preschool PATHS (Promoting Alternative THinking Strategies) Curriculum (Domitrovich, Greenberg, Kusche, & Cortes, 2005), which included brief social-emotional lessons, hands-on extension activities, and generalized teaching strategies (such as emotion coaching, social problem-solving dialogue) to promote child prosocial interaction and self-control. To promote emergent literacy skills, REDI used three curriculum components -- dialogic reading, sound games, and alphabet center activities. These components required teachers to lead an interactive reading activity daily, conduct brief sound games three times per week, and monitor children's skill

acquisition at an enriched alphabet center. Teachers were also trained to use enriched language (e.g., questions, reflections, rich vocabulary, and decontextualized talk) throughout the day (for more information on the study design and child outcomes, see Bierman et al., 2008).

During the implementation year, teachers received all necessary materials, participated in three days of training shortly before the school year began, as well as a one-day mid-year booster training. Teachers also received weekly mentoring visits from a REDI coach who spent 2–3 hours per week in the classroom and 1 hour per week in a meeting outside of the classroom, modeling appropriate techniques and providing technical assistance and support in implementing the curriculum (for a more detailed description of the intervention support model and teacher outcomes see Domitrovich et al., 2009). During the initial implementation year, teachers were compensated financially for the time spent meeting with the REDI coach, which occurred outside of the classroom as an add-on to their teaching contract.

Once the trial ended, during the year of the follow-up assessment, teachers were encouraged by the Head Start programs to continue their use of the REDI curriculum. REDI coaches made contact with teachers (by phone, email, or classroom visits) approximately once per month to check-in and answer any questions that teachers had, but teachers no longer had formal meetings with the coach nor any financial compensation for program preparation time.

Measures

Readiness for program implementation—The *organizational climate* of the workplace was assessed by teacher report, using a 13-item scale describing satisfaction with salary and benefits, quality of supervision, and clarity of roles, responsibilities, and communication in their Head Start program (Gill, Greenberg, Moon, & Margraf, 2007; $\alpha = .95$). *Teacher education levels*, as reported by teachers, were coded into four categories, indicating high school (1), associate's degree (2), four-year college degree (3), or master's degree (4). *Pre-intervention teaching quality* was assessed using the classroom observation systems described below.

Teaching quality—Teaching quality was assessed in the same way at the pre-intervention and follow-up assessments. At each time point, observation measures included the *Classroom Assessment Scoring System* (CLASS; La Paro, Pianta, & Stuhlman, 2004) which assessed 10 dimensions of teaching quality, each rated on a 7-point Likert scale after an observation period of 20 consecutive minutes. At each time point, four 20-minute epochs were rated in each classroom (usually on the same day). Ratings for each item were then averaged across the four epochs. Intraclass correlations reflecting the inter-rater reliability in the field ranged from ICC = .60 to ICC = .95 across the 10 items, median ICC = .77. As recommended by the scale developers when these data were collected, the ten behavioral rating items were summarized in two scores reflecting the domains of *Emotional Support* (positive climate, negative climate -reverse coded, teachers' sensitivity, over control-reverse coded, and behavior management, $\alpha = .86$) and *Instructional Support* (productivity, concept development, instructional learning formats, quality of feedback, $\alpha = .76$). At the same time, observers also completed the *Teaching Style Rating Scale* (TSRS; Domitrovich, Cortes, & Greenberg, 2000). Whereas the CLASS assessed the general quality of teacher-child interaction in the classroom, the TSRS was complementary because it rated the behavior of each teacher separately. It also focused more specifically on the teaching qualities that were emphasized in the REDI intervention. The TSRS included 9 items, each rated on a 5-point scale. During the course of data collection, average inter-rater agreement on TSRS items

was 93% (within 1 scale point). Intra-class correlations for individual items ranged from ICC = .60 to ICC = .75, median ICC = .68. Items in the TSRS were summed into three general scores: *Positive Discipline* (proactive/preventive approaches, use of praise, reinforcement and redirection, and absence of negative discipline domain, $r = .84$); *Classroom Management* (preparedness, use of consistent routines, and effective control and limit-setting, $r = .82$); and *Positive Emotional Climate* (emotion expression, support for student emotion regulation, and emotion modeling, $r = .71$).

A second coder visited classrooms at the same time and coded teacher language use with the *Classroom Language and Literacy Environment Observation* (CLEO; Holland-Coviello, 2005). On this measure, all child-directed utterances were classified into three categories, each represented as rates per minute. The categories and intraclass correlations reflecting their inter-rater reliability were: *directives*, reflecting commands and control statements (ICC = .96), *questions*, which elicited information from the child (ICC = .95), and *statements* which included all other neutral comments, usually providing information or commenting on ongoing activities (ICC = .92). Within each classroom context, observers also noted instances of *decontextualized talk* defined as utterances about people, places, or things not present (ICC = .97). In addition, immediately after each observation segment, observers completed a set of eight 5-point ratings describing the richness and sensitivity of teacher language use (e.g., vocabulary use, elaboration, cognitive challenge, responsiveness to child initiations, $r = .74$).

Prior to collecting data, separate training sessions were conducted by the CLASS developers, the TSRS developers, and the CLEO developers. In each case, research assistants were required to reach a criterion (80% agreement or higher, within one scale point) with training videotapes before moving into the field, and then were checked for observer drift over time (with inter-rater reliability collected on 15–20% of the observations.)

REDI program implementation quality—During the year of the research trial, when the REDI program was first introduced and supported in intervention classrooms, REDI coaches completed implementation ratings on a monthly basis to assess the quality with which teachers implemented the curriculum. After completing a set of specific ratings describing the fidelity, generalization, and child reactions, coaches rated the overall implementation quality for each curriculum component, using a 7-point Likert scale (1 = poor; 7 = exemplary), and this item was used in the present study analyses. The fidelity with which each of the 4 REDI curriculum components was implemented was rated separately each month; these monthly scores were then averaged to create a single score characterizing the implementation quality of that component for the academic year (stability over time was high for both components, ranging from $r = .93$ to $r = .82$). Ratings for dialogic reading, sound games, and alphabet center were averaged to create a composite score representing the overall implementation of the REDI language and literacy program. On average, teachers achieved ratings between 4 (adequate) to 5 (strong) for implementation quality during the year of the research trial, with a mean of 4.63 for PATHS and 4.53 for Literacy. (See Domitrovich et al., 2010 for more information about implementation.)

The same coach rating measure was used to assess REDI curriculum implementation during the year of follow-up assessment. In the spring of the year following the research trial, coaches completed ratings to describe the extent and quality of sustainability of the REDI program. Their ratings were based upon a classroom observation, review of teacher logs, and monthly check-ins with teachers. First, coaches completed a set of specific ratings describing the fidelity and generalization of different program elements and then gave each REDI curriculum component an overall implementation quality score on a 7-point scale (1 =

poor; 7 = exemplary). This item was used in the present study analyses. Sustainability of the literacy program components was represented by the averaged ratings for dialogic reading, sound games, and alphabet center ($r = .84$).

Teacher perceptions of the intervention—At the end of the research implementation year, teachers completed 13 items developed for this study describing their perceptions of the intervention. Ratings utilized a 5-point Likert scale (1 = virtually none to 5 = a great deal) and included the acceptability of the curriculum (e.g., user-friendly, easy to fit into their day), perceived effectiveness (e.g., engaging to child; positive impact), and program fit (e.g., matched their teaching style; was valued by their Head Start administrators). Teachers completed these ratings for each of the three literacy components, and these ratings were averaged to create a *Perceptions of Literacy* intervention score ($r = .93$). Teachers completed the 13 item rating on three components of the PATHS curriculum, which were averaged to create a *Perceptions of PATHS* intervention score ($r = .95$).

Data Collection Procedures

In the spring prior to the year of REDI research implementation and prior to randomization to intervention or control condition, research assistants visited classrooms and conducted baseline observations of teaching quality. In the fall, teachers reported on their education, experience, and workplace perceptions. During the year of the REDI research trial, coaches completed monthly ratings describing the quality of REDI implementation in the intervention classrooms. At the end of this research trial year, teachers in the intervention classrooms completed rating forms describing their perceptions of the intervention. Teachers were compensated financially for completing these ratings. During the subsequent year, REDI coaches checked in with teachers monthly (by phone, email, or in person), asked about program implementation and answered any questions teachers had. In the Spring, they conducted a classroom observation, reviewed teacher logs, and completed ratings of the quality of sustained REDI intervention implementation. In January-March of that follow-up year (two years after the base-line assessments of teaching quality), observers who were naïve concerning the intervention vs. control condition of the teachers, conducted observations to rate general teaching quality.

Results

Preliminary analyses were undertaken to compare the pre-intervention scores of the teachers who remained in the sample ($N = 37$) with those who left Head Start ($N = 7$) on the measures of teacher education, experience, and teaching quality used in this study. No significant differences emerged (all p 's $> .10$). Correlations among the measures collected at the follow-up assessment, including measures of sustained intervention quality and general teaching quality, are shown in Table 1. Coach ratings of the quality of sustained implementation of PATHS and the language/literacy intervention components were highly correlated, $r = .82$, $p < .001$, but in general these ratings were not significantly correlated with the observations of general teaching quality. The exceptions were that, at follow-up, classroom emotional climate was significantly correlated with PATHS curriculum implementation, and classroom emotional climate, behavior management, and positive discipline were all marginally correlated with the implementation of the language and literacy components. These associations suggest that teachers who maintained the intervention with high quality in one domain (social-emotional) also tended to sustain the intervention with high quality in the other domain (language-literacy), but this was not highly dependent upon teacher's concurrent general teaching skills.

Also evident in these correlations is that on average, associations among the different teaching practices measured in this study were positive and moderate in value, although a few measures showed strong associations (e.g., emotional support on the CLASS, classroom management and positive discipline on the TSRS each had $r > .70$). These findings suggest that the indices of teaching quality studied here tended to co-vary, but also represented semi-distinct domains of teaching quality.

Sustainability of the Social-Emotional and Language-Literacy Curriculum Components

Focusing only on the intervention classrooms, we tested the hypothesis that intervention teachers would continue to implement the social-emotional (PATHS Curriculum) and language-literacy (dialogic reading, sound games, alphabet center) curriculum components at rates equivalent to those exhibited during the research trial. A comparison of the means revealed consistency in the implementation quality of PATHS during the year of the research trial ($M = 4.63$, $SD = .71$) and the year of the follow-up assessment ($M = 4.55$, $SD = .90$). A paired t-test confirmed no significant differences in the implementation quality across the two time points, $t(19) = .47$, $p > .10$. In contrast, the mean level of sustainability for the literacy components of REDI in the year following the research trial was $M = 3.94$, $SD = .49$. A paired t-test revealed that this level of implementation quality was significantly lower than the implementation quality observed during the year of the research trial, $M = 4.53$, $SD = .51$, $t(19) = 3.58$, $p < .05$.

Factors Associated with Sustained High-Quality Use of the REDI Curriculum

Next, correlations were computed to assess links between factors that might be associated with sustained intervention quality, including both pre-intervention factors (e.g., organizational climate, teacher education levels, pretreatment teaching quality), and initial intervention experiences (e.g., teacher evaluations of the intervention, quality of initial intervention implementation). These correlations are displayed in Table 2. Note that these correlations must be interpreted with caution, given the small sample size of intervention classrooms ($N = 20$).

Pre-intervention organizational climate and teacher education showed small-to-moderate levels of association with sustained implementation quality, but these were not statistically significant in this small sample. Correlations linking pretreatment teaching quality with the sustained implementation of PATHS averaged $r = .47$, and all but one were statistically significant or marginally significant. Correlations linking pretreatment teaching quality with the sustained implementation of the language and literacy intervention components averaged $r = .35$, and five out of nine were statistically significant or marginally significant. These findings suggest that teachers who had stronger teaching skills at pre-intervention were more likely to sustain PATHS with high quality, and also tended to sustain the language and literacy components with high quality, relative to teachers with weaker pre-intervention skills. In terms of experiences with the intervention, teacher evaluations of both PATHS and the language-literacy intervention components were quite high, with means of 4.31 (S.D. = .49) and 4.19 (S.D. = .49) respectively, on a 5-point scale. These evaluations were only mildly and non-significantly correlated with sustained implementation quality, perhaps due to ceiling effects. Finally, the implementation quality of both PATHS and the literacy program components during their inaugural year were significantly or marginally correlated with program sustainability in the following year, ranging from $r = .59$, $p < .01$ to $r = .42$, $p = .06$, suggesting that teachers who learned and implemented REDI with greater fidelity during its initial introduction in the research trial were more likely than their peers to sustain the REDI curriculum components with high quality at the follow-up assessment one year later.

Intervention vs. Control Comparisons in Teaching Quality at the Follow-Up Assessment

Finally, we tested the hypothesis that the intervention would produce sustained improvements in general teaching quality that would be reflected in significant differences between intervention and control teachers at the follow-up assessment, in areas of both social-emotional support and language-literacy support. Means and standard deviations for teachers in the intervention and control conditions, along with the results of multiple regression models that controlled for cohort, setting (rural vs. urban), and tested the impact of intervention condition (0 = control, 1 = intervention) are presented in Tables 3 and 4. Estimates of effect size (d) were derived from the regression models and were calculated as the difference in the adjusted means of the intervention and control groups divided by the pooled standard deviation (Cohen, 1988).

Consider first the impact on teaching quality in the area of social-emotional support (Table 3). Results are shown for the summary scale scores (underlined in the table) and then for each of the contributing rating items to allow for comparisons with the post-treatment results (see Domitrovich et al., 2009). Significant group differences favoring the intervention group emerged on the Emotional Climate scale of the TSRS, $d = 1.12$, $p < .01$, supported by significant or marginally significant group differences on the contributing items evaluating teacher emotional expression, emotion regulation, and emotional modeling. Significant group differences favoring the intervention group also emerged on the Emotional Support scale of the CLASS, $d = .72$, $p < .01$, supported by significant intervention group elevations on the positive climate rating and marginally significant reductions on the negative climate rating. Marginally-significant elevations for the intervention group were also evident on the Classroom Management scale of the TSRS, $d = .56$, $p < .10$, supported by significant group differences on the preparedness item rating. Finally, significant group differences favoring the intervention group emerged on the Positive Discipline scale of the TSRS, $d = .94$, $p < .01$, supported by significant intervention elevations in proactive/preventive strategies and the absence of negative discipline. As discussed further below, these effects were directly comparable to the impact of intervention on teaching quality documented at post-treatment.

Next, consider the impact on teaching quality in the domain of cognitive-linguistic support (Table 4). In the area of language use, significant group differences favoring the intervention emerged on the rate of questions asked in the classroom, $d = .97$, $p < .01$, along with marginally-significant effects on the rate of statements, the use of decontextualized talk, and the ratings of the sensitivity and richness of talk. These effects are comparable, although slightly smaller, than at post-treatment. No significant group differences emerged on the Instructional Support scale of the CLASS, nor any of its contributing items, although means for the intervention group were consistently higher than those of the control group.

Discussion

Evidence-based preschool interventions can improve the school readiness of children from disadvantaged backgrounds, maximizing the value of early education programs (Burchinal et al., 2002; Ramey & Ramey, 2004). However, to have widespread impact, these interventions must be implemented with high quality and sustained over time (Elias, Zins, Graczyk, & Weissberg, 2003; Han & Weiss, 2005). Addressing the need to explore sustainability in the context of early childhood education, this study provides initial data on the degree to which evidence-based preschool curriculum components and more general aspects of teaching quality were sustained in the year after their initial introduction. Several important issues are identified for further research.

Sustainability of the REDI Components and Predictive Correlations

One unique feature of this study was its focus on the sustainability of curriculum components and teaching quality in two distinct domains – social-emotional programming (e.g., the PATHS Curriculum) and language-literacy programming (e.g., dialogic reading, sound games, alphabet center). A prior analysis of REDI sustainability that focused on the perspectives of the participating teachers suggested that the REDI social-emotional component (PATHS Curriculum) was sustained with better quality and more teacher enthusiasm than the emergent literacy intervention components (Sanford DeRousie & Bierman, 2012). This study extended those results by including quantitative analyses of coach ratings of curriculum implementation quality during the year of follow-up assessment and by examining their predictive correlates. In this study, coach ratings showed sustained high-quality implementation of the PATHS Curriculum, with mean levels above 4.5 during both the research implementation and follow-up assessments, indicating 4 (adequate) to 5 (strong) implementation quality. In contrast, coach ratings revealed a significant decline in the quality of implementation of the language-literacy intervention components, dropping from a mean of 4.55 during the research trial to a mean of 3.94 during the follow-up year, under the threshold of 4 (adequate).

Although absolute levels of sustained program use varied across the social-emotional and language-literacy domains, in general, teachers who sustained evidence-based programming more effectively in one domain also tended to do so in the other domain, reflected in significant correlations in the sustainability quality of PATHS and language-literacy components ($r = .82$). Exploratory correlations linking pre-intervention readiness, intervention experiences, and intervention sustainability also showed similar patterns across the two curriculum domains. Strong and statistically significant associations emerged between pre-intervention teaching quality (especially emotional support, classroom management, positive discipline, instructional support, and the use of questions) and sustained implementation of the curriculum components. Initial implementation quality during the inaugural year of intervention also emerged as a significant predictor of subsequent implementation quality.

These results suggest that teachers who, prior to the initiation of the intervention, were able to manage their classrooms effectively with warm support and positive discipline were more likely to sustain the REDI curriculum with high-fidelity than teachers with lower ratings in these areas. Pre-intervention levels of questions and instructional support (e.g., the pacing and presentation of learning activities and quality of instructional feedback) also predicted sustained intervention implementation two years later. In general, predictive correlations between observations of pre-intervention teaching quality and intervention sustainability two years later were stronger than concurrent correlations between intervention implementation and teaching quality during the follow-up year (which were largely non-significant). This pattern is consistent with the interpretation that initial teaching quality facilitated uptake and sustained high-quality use of the new curriculum components, and at the same time, the introduction of the curriculum and professional development facilitated improvements in general teaching quality over time, attenuating concurrent associations between teaching quality and quality of curriculum implementation by the end of the follow-up year.

There are several reasons why pre-intervention teaching quality might improve the uptake and sustained high-quality use of evidence-based intervention components. Teachers who were already more competent in providing social-emotional support and organizing instruction in their classrooms may have found that the REDI curriculum components were easier and more comfortable to implement, and therefore, found them easier and more comfortable to sustain. A teacher who is able to foster positive relationships with her

students may also be better prepared and better able to adjust to a new curriculum because she is not as distracted by classroom management challenges (Hughes et al., 2005).

The implications of these findings for practice are not straightforward. On the one hand, they suggest that efforts to improve the general teaching skills of teachers with low scores prior to introducing evidence-based curriculum might improve the long-term sustained quality with which that curriculum is implemented. On the other hand, the use of evidence-based curriculum appears to moderate the impact of poor teaching quality on child outcomes, such that teachers with lower teaching quality produce better child outcomes when they use evidence-based curriculum components (see Domitrovich, Bierman, Nix, Gill, & Gest, in press). In addition, experiences with evidence-based curriculum components can improve general teaching quality, when combined with professional development support as in REDI. For this reason, it is likely better to initiate the use of evidence-based curriculum components even when teachers have lower teaching quality, and adjust the level of on-going professional development support and coaching provided in order to help these teachers improve their teaching quality. One strategy to accomplish this may be to identify thresholds of effective teaching quality (Burchinal, Vandergrift, Pianta, & Mashburn, 2010) and orient professional development support to helping teachers attain those thresholds (e.g., criterion-referenced coaching) (Fox & Hemmeter, 2009). Another approach might be to use an adaptive approach to deliver professional development support (Collins, Murphy, & Bierman, 2004), varying the intensity of coaching during implementation in order to provide more intensive coaching to less skillful teachers, allowing them to bootstrap more general improvements in their teaching skills through their participation in the professional development associated with the implementation of the evidence-based intervention.

The significant correlations between implementation fidelity during the research trial and high-quality sustainability during the year following the trial supports the theoretical perspective that sustainability should be seen as a process that begins with the very first implementation activities, not just something tacked on at the end of the intervention process (Pluye et al, 2004; Scheirer, 2005). In a way, these results are particularly notable, given that rates of fidelity were quite high in general during the research trial, as a function of the extensive coaching and classroom support that teachers received from the REDI coaches. That there were significant associations even with this restricted range of scores suggests that it is especially important for schools looking to implement and sustain evidence based practices to pay close attention to insuring high quality implementation from the very initiation of the program if the program is to be sustained, with quality, over time.

Prior research suggested that teacher education and organizational climate might also predict sustained use of evidence-based programs (Domitrovich et al., 2010; Whitehurst & Lonigan, 1998). However, neither of these factors correlated significantly with sustained program use in this study. The size of the correlations found here between teacher education, organizational climate, and sustained implementation quality of the REDI curriculum components were of a similar magnitude to those documented in other studies (e.g., r s ranging from .25 to .33), but given the small sample size in this study, these associations were not statistically significant. Perceptions of the intervention showed only small and non-significant associations with sustained implementation quality, but these associations were likely highly attenuated by ceiling effects on the teacher evaluation scales. That is, most teachers found the intervention to be very engaging for children, comfortable to use, and effective in promoting children's skill development. Each of these factors (organizational climate, teacher education levels, and perceptions of the intervention) should receive additional attention in future research examining intervention program sustainability.

In terms of explaining the higher levels of sustainability of the PATHS Curriculum compared with the language and literacy components, this study offers little insight, as the patterns of prediction of sustainability were very similar across these two domains. In qualitative interviews with teachers, they noted that PATHS filled a void for them, as they had no other curriculum for social-emotional skills development, and they appreciated the “ready-made” lesson plans and activities (see Sanford DeRousie & Bierman, 2012 for details). In contrast, teachers were already reading with their children and teaching letters prior to the REDI intervention. In addition, enriching curriculum to promote emergent literacy skills had become a Head Start, state, and federal priority during the years immediately preceding this study. Hence, teachers were using and/or were introduced to a number of alternative approaches to promoting emergent literacy (e.g., shared reading, Kid Writing), and alternative curriculum materials (e.g., Alpha-tales) that overlapped in content and mission with the REDI literacy components. Teachers expressed concerns about fitting in all of the literacy activities and, understandably, many wanted to blend or choose among approaches (Sanford DeRousie & Bierman, 2012). Teaching efforts in the language and literacy area may thus have been dispersed across different approaches and curricula, contributing to a decline in REDI implementation. The current findings do not speak to this issue directly, but they are consistent with this interpretation, because they do not reveal any notable differences in the correlates of the sustained use of the language and literacy vs. social-emotional components. Further research comparing sustainability processes for different kinds of preschool interventions is needed. Because this study involved only one program, it is not possible to determine whether the findings related to the differential sustainability of social-emotional vs. language-literacy interventions will hold more generally, or whether they are specific to the REDI intervention components or the broader context of Head Start focus at the time of this study.

Intervention and Control Group Differences in Teaching Quality at Follow-Up

Intervention-control group differences in teaching quality at the follow-up assessment were equivalent to, or in some cases, larger than the differences documented at the post-test assessment. For example, in the domain of social-emotional support, post-treatment effects at post-treatment (the end of the first year of intervention implementation) showed medium effect sizes, in the $d = .40 - .60$ range. By the follow-up assessment one year later, effect sizes were in the medium to large range for teaching quality in areas of emotional support ($d = .72$), emotional climate ($d = 1.12$), positive discipline ($d = .94$), and classroom management ($d = .56$). In the area of cognitive support and language use in the classroom, a moderate effect size was maintained on CLASS instructional support ($d = .48$ at post-test, $d = .48$ at follow-up), although this effect was not statistically significant in this small sample. Effects on language use at post-test were moderate to large, ranging from $d = .67 - .89$, and were mostly in the moderate range at follow-up ($d = .49 - .56$), with the exception of questions, which maintained a large effect ($d = .97$). Overall, these findings suggest that gains in general teaching strategies that were associated with the intervention were well-maintained after withdrawal of the intensive professional development support. REDI provided a substantial level of professional development support to teachers, including four days of workshop training, and weekly classroom visits and mentoring meetings. It is heartening to see that this investment produced sustained improvements in general teaching quality. However, many programs lack the resources to duplicate this level of professional development support for new curriculum components. This raises an important question for future research: can more cost-effective methods of professional development delivery be equally effective? Initial evaluations of technology-facilitated coaching appear promising in this regard (Pianta, Mashburn, Downer, Hamre, & Justice, 2008).

Limitations and Future Directions

There were several limitations to this study. First, the sample size was quite small. Only relatively large effects ($d > .65$) were detectable at the statistically significant level with this sample size. A larger sample may have allowed for the documentation of significant impact for features that had smaller effect sizes but were nonetheless influences on sustainability. The small sample size (20 intervention and 17 control teachers) also made the use of multiple imputation inadvisable, as it is generally recommended that for each parameter in the imputation model, there are 10 variables in the data set. Hence, we were not able to include the 7 teachers who left the study at the end of the intervention year which otherwise would have increased power somewhat. Second, following teachers beyond the first year of sustainability would have added depth to the study. Since sustainability is often considered to be the extent to which a program is sustained for at least two years after the implementation year (Elias, 2004), focusing on the first year after initial implementation captured only the initial transition to sustainability. A third limitation is that this study does not include all of the potential factors that may influence sustainability. Factors such as the quality/quantity of training, teacher understanding of the intervention theory, teacher turnover, and the quality of on-going technical assistance, as well as broader factors such as the political climate and the educational structure of the school system, may all influence sustainability (Florian, 2001; Greenberg et al., 2005). Fourth, the ratings of fidelity of implementation collected during the research trial and sustainability year were provided by the same REDI coaches, creating a possibility of same-rater inflation. However, the observations of teaching quality at the pre-intervention and follow-up assessments were conducted by a completely different set of observers who were naive concerning the intervention vs. control status of the teachers, thus providing unbiased measures. Fifth, it is important to note that the within-intervention group analyses step away from the experimental design, and therefore may be confounded by selection factors that account for significant associations. Finally, it is unclear how much the monthly coach check-ins facilitated program sustainability. In general, evidence-based programs may be sustained more effectively if an on-going system is in place which both monitors sustained intervention implementation and provides technical assistance, as needed. Future research may be helpful in clarifying the best way to organize these kinds of systematic efforts to monitor and increase the sustained, high-fidelity use of evidence-based curriculum components in preschool settings.

The noted limitations reduce the degree to which firm conclusions can be drawn about the generalizability of the findings. However, this preliminary study validates the importance of including follow-up assessments of teachers and attending to processes of sustainability to further the goal of promoting long-term improvements in preschool educational programs and practices. The results suggest that sustainability requires more focused attention in program planning in order to promote the high-quality sustained use of evidence-based curriculum components and teaching practices that are needed to substantially improve learning outcomes for children. Promoting the sustainability of school-based evidence-based programs is important, as it allows multiple cohorts of children to benefit from the same program, provides consistency for teachers, and is cost-effective for school systems. In the long run, the quality of preschool programming and needs of children will be best served by strategies that facilitate the institutionalization of evidence-based practices, so that they become embedded effectively and permanently within the culture of preschools.

References

- Bierman, K.L.; Domitrovich, C.; Darling, H. Early Prevention Initiatives. In: Roopnarine, J.; Johnson, J., editors. *Approaches to Early Childhood Education*. 5th Ed.. Columbus, OH: Pearson Merrill Prentice Hall; 2009. p. 147-164.

- Bierman KL, Domitrovich CE, Nix RL, Gest SD, Welsh J, Greenberg MT, Gill S. Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program. *Child Development*. 2008; 79:1802–1817. [PubMed: 19037951]
- Blau DM. The production of quality in child-care centers: Another look. *Applied Developmental Science*. 2000; 4:136–148.
- Burchinal MR, Peisner-Feinberg ES, Pianta R, Howes C. Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology*. 2002; 40:415–436.
- Burchinal MR, Vandergrift N, Pianta R, Mashburn A. Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten program. *Early Childhood Research Quarterly*. 2010; 25:166–176.
- Clarke D, Hollingsworth H. Elaborating a model of teacher professional growth. *Teaching and Teacher Education*. 2002; 18:947–967.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd ed.. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Collins LM, Murphy SA, Bierman KL. A conceptual framework for adaptive preventive interventions. *Prevention Science*. 2004; 5:185–196. [PubMed: 15470938]
- Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: Are implementation effects out of control. *Clinical Psychology Review*. 1998; 18:23–45. [PubMed: 9455622]
- de Kruif RE, McWilliam RA, Ridley SM, Wakely MB. Classification of teachers' interaction behaviors in early childhood classrooms. *Early Childhood Research Quarterly*. 2000; 15:247–268.
- Domitrovich, CE.; Bierman, KL.; Nix, RL.; Gill, S.; Gest, SD. Improving preschool education with curriculum enhancements and professional development: The REDI intervention model. In: Howes, C.; Hamre, BK.; Pianta, RC., editors. *Effective professional development in early childhood education*. Baltimore, MD: Brookes; in press
- Domitrovich, CE.; Cortes, R.; Greenberg, MT. The Teacher Style Rating Scale technical report. Pennsylvania State University; 2000. Unpublished manuscript
- Domitrovich CE, Gest SD, Gill S, Bierman KL, Welsh J, Jones D. Fostering high quality teaching in Head Start classrooms: Experimental evaluation of an integrated enrichment program. *American Education Research Journal*. 2009; 46:567–597.
- Domitrovich CE, Gest SD, Jones J, Gill S, Sanford DeRousie R. Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly*. 2010; 25(3):284–298. [PubMed: 22844183]
- Domitrovich, CE.; Greenberg, MT.; Kusche, C.; Cortes, R. *The Preschool PATHS Curriculum*. Channing Bete Publishing Company; 2005.
- Durlak JA, DuPree EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*. 2008; 41:327–350. [PubMed: 18322790]
- Dusenbury L, Brannigan R, Hansen WB, Walsh J, Falco M. Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research*. 2005; 20:308–313. [PubMed: 15522898]
- Elias, ME. Sustaining social-emotional learning programs: A study of the developmental course of model/flagship SEL sites. 2004. Unpublished study
- Elias M, Zins JE, Graczyk PA, Weissberg RP. Implementation, sustainability and scaling-up of social-emotional and academic innovations in public schools. *School Psychology Review*. 2003; 32(3): 303–319.
- Elliott DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. *Prevention Science*. 2004; 5:47–53. [PubMed: 15058912]
- FACES. Head Start Teacher Self-Administrated Survey: Head Start Family and Children Experiences Survey (FACES). 1999.
- Florian, J. Sustaining education reform: Influential factors. Aurora, CO: Mid-Continent Research for Education and Learning. (ERIC Document Reproduction Service No. ED4535833); 2001.

- Fox, L.; Hemmeter, ML. A program-wide model for supporting social emotional development and addressing challenging behavior in early childhood settings. In: Sailor, W.; Dunlap, G.; Sugai, G.; Horner, R., editors. *Handbook of Positive Behavior Support*. New York: Springer; 2009. p. 177-202.
- Gill S, Greenberg MT, Moon C, Margraf P. Home visitor competence, burnout, support, and client engagement. *Journal of Human Development in the Social Environment*. 2007; 15(1):23-44.
- Gottfredson GD, Jones EM, Gore TW. Implementation and evaluation of a cognitive-behavioral intervention to prevent problem behavior in a disorganized school. *Prevention Science*. 2002; 3:43-56. [PubMed: 12002558]
- Greenberg, MT.; Domitrovich, CE.; Graczyk, PA.; Zins, JE. *The study of implementation in school-based preventive interventions: Theory, research, and practice*. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration; 2005.
- Hall, GE.; Hord, SM. *Implementing change: Patterns, principles, and potholes*. Boston: Allyn and Bacon; 2001.
- Han S, Weiss B. Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*. 2005; 33:665-679. [PubMed: 16328743]
- Hughes JN, Cavell TA, Meehan BT, Zhang D, Collie C. Adverse school context moderates the outcomes of selective interventions for aggressive children. *Journal of Consulting and Clinical Psychology*. 2005; 73:731-736. [PubMed: 16173861]
- Joseph GE, Strain PS. Comprehensive evidence-based social-emotional curricula for young children: An analysis of efficacious adoption potential. *Topics in Early Childhood Special Education*. 2003; 23:65-76.
- Klinger JK, Vaughn S, Hughes MT, Arguelles ME. Sustaining research-based practices in reading: A three year follow-up. *Remedial and Special Education*. 1999; 20:263-274.
- LaParo KM, Pianta RC, Stuhlman M. The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*. 2004; 104:409-426.
- Lonigan CJ. Development, assessment, and promotion of pre-literacy skills. *Early Education and Development*. 2006; 17:91-114.
- Pianta RC, Mashburn AJ, Downer JT, Hamre BK, Justice LM. Pianta. Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*. 2008; 23:431-451.
- Pluye P, Potvin L, Denis JL. Making public health programs last: Conceptualizing sustainability. *Evaluation and Program Planning*. 2004; 27:121-133.
- Ramey CT, Ramey SL. Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly*. 2004; 50:471-491.
- Sanford DeRousie RM, Bierman KL. Examining the sustainability of an evidence-based preschool curriculum: The REDI program. *Early Childhood Research Quarterly*. 2012; 27:55-65. [PubMed: 22408287]
- Sarama, J.; Clements, DH. *Early childhood mathematics education research: Learning trajectories for young children*. New York: Routledge; 2009.
- Scheirer MA. Is sustainability possible? A review and commentary on empirical studies of program sustainability. *The American Journal of Evaluation*. 2005; 26:320-347.
- Vaughn S, Klingner J, Hughes M. Sustainability of research-based practices. *Exceptional Children*. 2000; 66:163-171.
- Weaver RH. Predictors of quality and commitment in family child care: Provider education, personal resources, and support. *Early Education & Development*. 2002; 13:265-282.
- Whitehurst GJ, Lonigan CJ. Child development and emergent literacy. *Child Development*. 1998; 69:848-872. [PubMed: 9680688]

Table 1

Correlations Among Measures Collected at the Follow-Up Assessment

	2	3	4	5	6	7	8	9	10	11
<u>Intervention Quality</u>										
1. PATHS	.82**	.50*	.14	.07	.22	.08	.13	-.20	-.33	.21
2. Literacy	--	.42 ⁺	.20	.41 ⁺	.39 ⁺	.13	.19	-.08	-.29	.28
<u>Teaching Quality</u>										
3. Emotional climate ^a	--	.35*	.38*	.43**	.59**	.28 ⁺	.17	.18	.48**	
4. Emotional support ^b	--	--	.86**	.75**	.65**	.29 ⁺	.45*	.30 ⁺	.58**	
5. Class management ^a	--	--	.75**	.70**	.35*	.43**	.29 ⁺	.65**		
6. Positive discipline ^a	--	--	.49**	.47**	.42**	.45**	.70**			
7. Instructional support ^b	--	--	.40*	.31 ⁺	.13	.48**				
8. Statements ^c	--	--	.39*	-.01	.55**					
9. Questions ^c	--	--	.46**	.59**						
10. Decontextualized ^c	--	--	.41*							
11. Rich-sensitive talk ^c	--	--								

Note: Sample size is N=20 for correlations with intervention quality (intervention classrooms only) and N=37 for teaching quality (intervention and control classrooms)

^aTSRS

^bCLASS

^cCLEO

⁺ $p < .10$.

* $p < .05$.

** $p < .01$.

Table 2

Correlations Linking Pre-Intervention Readiness and Initial Intervention Experiences with Sustained Intervention Implementation

Factors Associated with Sustained Intervention Quality	Sustained Curriculum Implementation Quality	
	PATHS	Literacy
<u>Pre-intervention Readiness</u>		
Organizational climate	.28	.25
Teacher education	.26	.33
Pre-intervention Emotional Climate TSRS	.40**	.25
Pre-intervention Emotional support CLASS	.55*	.41+
Pre-intervention Class management TSRS	.61**	.44*
Pre-intervention Positive Discipline TSRS	.59**	.48*
Pre-intervention Instructional Support CLASS	.57**	.52*
Pre-intervention Statements CLEO	.23	.03
Pre-intervention Questions CLEO	.49*	.53*
Pre-intervention Decontextualized CLEO	.42+	.28
Pre-intervention Language Richness CLEO	.41+	.25
<u>Intervention Experiences</u>		
Teacher Evaluation of PATHS	.16	.19
Teacher Evaluation of Literacy Components	-.05	.16
Initial Implementation Quality - PATHS	.54**	.59**
Initial Implementation Quality - Literacy	.42+	.49*

Table 3

Sustained Effects on Teaching Quality: Social-Emotional Support

Observed Teaching Quality	Intervention		Control		Estimate	S.E.	d
	Mean	(S.D.)	Mean	(S.D.)			
<u>Emotional Climate TSRS</u>	3.77	(1.07)	2.65	(0.91)	1.13**	(0.34)	1.12
Emotion expression	4.15	(1.18)	2.94	(1.09)	1.22**	(0.38)	1.06
Emotion regulation	3.45	(1.39)	2.71	(0.98)	.75 ⁺	(0.42)	.61
Emotion modeling	3.70	(1.17)	2.29	(0.98)	1.42**	(0.37)	1.29
<u>Emotional support CLASS</u>	5.56	(0.49)	5.16	(0.63)	.40*	(0.19)	.72
Positive climate	6.13	(0.70)	5.51	(0.84)	.62*	(0.26)	.81
Teacher sensitivity	5.58	(1.00)	5.06	(0.86)	.53	(0.32)	.55
Overcontrol (R)	6.56	(0.50)	6.25	(0.77)	.32	(.20)	.49
Negative climate (R)	6.82	(0.32)	6.56	(0.55)	.27 ⁺	(0.15)	.61
Behavior management	5.71	(0.79)	5.43	(0.84)	.29	(0.28)	.35
<u>Class management TSRS</u>	4.53	(0.64)	4.18	(0.62)	.37 ⁺	(0.21)	.56
Control / Limit-setting	4.55	(0.76)	4.17	(0.81)	.40	(0.25)	.48
Preparedness	4.45	(0.83)	3.88	(0.78)	.57*	(0.27)	.70
Consistency / Routine	4.60	(0.68)	4.47	(0.62)	.13	(0.23)	.20
<u>Positive Discipline TSRS</u>	4.42	(0.59)	3.76	(0.80)	.67**	(0.21)	.94
Proactive/preventive	4.25	(0.91)	3.64	(0.86)	.63*	(0.28)	.68
Behavior management	4.35	(0.93)	3.94	(0.83)	.43	(0.29)	.46
Absence of negative	4.65	(0.93)	3.71	(1.36)	.95**	(0.34)	.82

Note: Negative climate is reverse-scored.

⁺ $p < .10$.

* $p < .05$.

** $p < .01$.

Table 4

Sustained Effects on Teaching Quality: Cognitive-Linguistic Support

	Intervention		Control		Estimate	S.E.	<i>d</i>
	Mean	(S.D.)	Mean	(S.D.)			
<u>Instructional Support CLASS</u>	4.05	(0.87)	3.51	(0.52)	.38	(0.24)	.48
Productivity	5.11	(1.01)	4.88	(0.83)	.49	(0.32)	.49
Quality of feedback	3.67	(1.00)	2.87	(0.81)	.42	(0.33)	.38
Concept development	3.24	(1.05)	2.74	(0.78)	.30	(0.27)	.33
Instructional formats	4.15	(1.15)	3.54	(0.63)	.31	(0.38)	.28
<u>Linguistic Support CLEO</u>							
Statements	1.57	(0.32)	1.55	(0.35)	.20 ⁺	(0.12)	.49
Questions	1.77	(0.33)	1.71	(0.52)	.60 ^{**}	(0.20)	.97
Decontextualized talk	.52	(0.56)	.44	(0.48)	.45 ⁺	(0.26)	.56
Richness of talk	2.66	(0.68)	2.37	(0.53)	.31 ⁺	(0.16)	.56

⁺ $p < .10$.

* $p < .05$.

** $p < .01$.

Table 5

Correlations Linking Predictors to Follow-Up Assessment

Factors Associated with Sustained Intervention Quality	Sustained Curriculum Implementation Quality	
	PATHS	Literacy
<u>Pre-intervention Readiness</u>		
Organizational climate	.28	.25
Teacher education	.26	.33
Pre-intervention Emotional Climate TSRS	.40 ⁺	.25
Pre-intervention Emotional support CLASS	.55 [*]	.41 ⁺
Pre-intervention Class management TSRS	.61 ^{**}	.44 [*]
Pre-intervention Positive Discipline TSRS	.59 ^{**}	.48 [*]
Pre-intervention Instructional Support CLASS	.57 ^{**}	.52 [*]
Pre-intervention Statements CLEO	.23	.03
Pre-intervention Questions CLEO	.49 [*]	.53 [*]
Pre-intervention Decontextualized CLEO	.42 ⁺	.28
Pre-intervention Language Richness CLEO	.41 ⁺	.25
<u>Intervention Experiences</u>		
Teacher Evaluation of PATHS	.16	.19
Teacher Evaluation of Literacy Components	-.05	.16
Initial Implementation Quality - PATHS	.54 ^{**}	.59 ^{**}
Initial Implementation Quality - Literacy	.42 ⁺	.49 [*]

Note: Note: Correlations are for intervention teachers ($N = 20$).

⁺ $p < .10$.

^{*} $p < .05$.

^{**} $p < .01$.