HHS Public Access

Author manuscript

Prev Sci. Author manuscript; available in PMC 2020 February 10.

Published in final edited form as:

Prev Sci. 2018 February ; 19(2): 186–196. doi:10.1007/s11121-017-0832-y.

Factors Associated with Teacher Delivery of a Classroom-based Tier 2 Prevention Program

Kevin S. Sutherland[1], **Maureen A. Conroy**[2], **Bryce D. McLeod**[1], **James Algina**[2], **Rachel L. Kunemund**[1]

[1]Virginia Commonwealth University

[2]University of Florida

Abstract

Teachers sometimes struggle to deliver evidence-based programs designed to prevent and ameliorate chronic problem behaviors of young children with integrity. Identifying factors associated with variations in the quantity and quality of delivery is thus an important goal for the field. This study investigated factors associated with teacher treatment integrity of BEST in CLASS, a tier-2 prevention program designed for young children at risk for developing emotional/behavioral disorders. Ninety-two early childhood teachers and 231 young children at-risk for emotional/behavioral disorders participated in the study. Latent growth curve analyses indicated that both adherence and competence of delivery increased across six observed time points. Results suggest that teacher education and initial levels of classroom quality may be important factors to consider when teachers deliver tier-2 (i.e., targeted to children who are not responsive to universal or tier-1 programming) prevention programs in early childhood settings. Teachers with higher levels of education delivered the program with more adherence and competence initially. Teachers with higher initial scores on the Emotional Support subscale of the Classroom Assessment Scoring System (CLASS) delivered the program with more competence initially and exhibited higher growth in both adherence and competence of delivery across time. Teachers with higher initial scores on the Classroom Organization subscale of the CLASS exhibited lower growth in adherence across time. Contrary to hypotheses, teacher self-efficacy did not predict adherence, and teachers who reported higher initial levels of student engagement self-efficacy exhibited lower growth in competence of delivery. Results are discussed in relation to teacher delivery of evidence-based programs in early childhood classrooms.

Keywords

early childhood; teacher delivery; problem behavior

[1]Two items assess Rules: *3–5 rules are visible in classroom* and *Teacher reviews rules, addresses rule violations*. Only the second item was included in the current study as we focused on teacher-delivered practices; *3–5 rules are visible in classroom* is a static item that did not change from observation to observation.

Compliance with Ethical Standards

Ethical approval. All study procedures involving human participants were in accordance with the ethical standards of the researchers' Institutional Review Boards and with the 1964 Helsinki declaration and its later amendments of comparable ethical standards.

Informed consent. Informed consent was obtained from all individual participants included in the study.

Disclosure of potential conflicts of interest. The authors declare that they have no conflict of interest.

Factors Associated with Teacher Implementation of a Classroom-based Tier 2 Prevention Program

With the national emphasis on early childhood education and the expansion of federal and state funded early childhood programs, practitioners have seen an increase in the number of young children who enter these programs and demonstrate chronic problem behavior that can impact their developmental outcomes (Carter et al., 2010; McCabe & Altamura, 2011). Early childhood teachers struggle with providing high quality instruction for these young children, who often present with multiple and cumulative risk factors (e.g., exposure to harsh parenting practices, living in poverty and violent communities; Berlin, Brooks-Gunn, McCarton & McCormick, 1998; Nelson, Stage, Duppong-Hurley, Synhorst & Epstein, 2007). Evidence-based programs (EBPs) that target amelioration of young children's social, emotional, and behavioral problems do exist (e.g., see *PK - Promoting Alternative Thinking Strategies [PK-PATHS]*, Domitrovich, Cortes, & Greenberg, 2007; *Incredible Years*, Webster-Stratton, Reid, & Hammon, 2004), but the field has struggled to demonstrate their widespread effectiveness and sustainability (Domitrovich, Gest, Jones, Gill & DeRousie, 2010; Durlak, 2010).

Implementation science seeks to address this challenge by attempting to understand how EBPs can be made more effective and sustainable in authentic community-based early childhood settings (Fixsen et al., 2005). Applied to EBPs targeted at young children who demonstrate chronic problem behavior that places them at risk for emotional/behavioral disorders (EBD) in early childhood settings, implementation science focuses on transferring efficacious EBPs from research settings into authentic classroom settings. However, delivery of these EBPs in authentic early childhood settings by teachers can be difficult due to the complexity of the programs and the contexts in which they are implemented (Durlak, 2010, 2015). Thus, an important focus of implementation research is to identify factors that influence the delivery of EBPs across a variety of settings (Mendel, Meredith, Schoenbaum, Sherbourne, & Wells, 2008; Proctor et al., 2011).

Treatment integrity, also referred to as treatment fidelity, fidelity of implementation, and intervention integrity (Dane & Schneider, 1998; McLeod et al., 2013; Sanetti & Kratochwill, 2009), is defined as the extent to which an intervention is delivered as intended. Two components of treatment integrity are considered important in implementation research: treatment adherence and competence of delivery (McLeod et al., 2013; Schoenwald et al., 2011). As these components relate to the delivery of EBPs by teachers in early childhood settings, treatment *adherence* represents the extent to which a teacher delivers the intervention program as designed while *competence* refers to the level of skill with which a teacher delivers program components. Each component assesses unique aspects of program delivery by teachers and is important to understanding how and why interventions may or may not be effective when delivered in authentic settings by teachers (Carroll & Nuro, 2002; McLeod et al., 2013).

To understand the factors that might influence the implementation of EBPs in school settings, researchers have used social-ecological frameworks (e.g., Bronfenbrenner, 1979) to

inform the development of conceptual models (e.g., Domitrovich et al., 2008; Han & Weiss, 2005). These models identify factors hypothesized to influence implementation at a variety of levels, including (a) macro-level factors (e.g., policies, funding); (b) school and program level factors (e.g., school culture and climate, administrative support and leadership); and (c) teacher level factors (e.g., training, experience). While researchers acknowledge the importance of examining how different levels of the models impact implementation (Domitrovich et al., 2008; Durlak, 2015; Han & Weiss, 2005), recent work suggests that broader factors (e.g., macro and school-level factors) may be less influential than more proximal (e.g., classroom and teacher factors) for interventions delivered by teachers (Domitrovich et al., 2015). Given these findings and the number of EBPs that utilize teachers in implementation of programs within classrooms, the importance of understanding both teacher and classroom factors associated with treatment integrity is highlighted, and is the area of focus for this paper. Specifically the current study investigated factors associated with teacher implementation integrity of BEST in CLASS (Conroy et al., 2015), a tier-2 prevention program that is designed to ameliorate young children's risk for developing EBD.

BEST in CLASS

Tier-2 programs are typically delivered to children who are not responsive to universal, tier-1 programming (Bruhn, Lane & Hirsch, 2014), and are often characterized by systematic screening of children at elevated risk for learning and behavioral difficulties and are delivered to small groups of children who exhibit similar problems (Mitchell, Bruhn, & Lewis, 2015). BEST in CLASS is a theoretically-grounded tier-2 intervention that focuses upon improving the quantity and quality of early childhood teachers' use of effective instructional practices with young children identified with chronic problem behaviors that place them at risk for EBD. Unlike many other tier-2 interventions that are delivered in small groups, teachers learn and are coached to implement BEST in CLASS practices specifically with the identified focal children in their classrooms during naturally occurring classroom activities. Teachers are taught to implement the BEST in CLASS model, which is comprised of key instructional practices that focus on increasing effective teacher-child interactions. This process also acknowledges the transactional nature of social interchanges (Sameroff, 1995) and how affecting behavior and transactions (e.g., improving teacher-child interactions) can influence the child's broader ecology (Bronfenbrenner, 2005), in this case the classroom environment.

BEST in CLASS is a manualized tier 2 intervention that includes a one-day training on the BEST in CLASS practices followed by 14 weeks of practice-based coaching to support teachers in implementation of the model (see Conroy et al., 2015 for a description of the BEST in CLASS intervention model). Previous investigations of BEST in CLASS have demonstrated promise on a variety of child outcomes, including significant reductions in disruptive behavior and negative teacher-child interactions and significant increases in engagement and positive teacher-child interactions (Conroy et al., 2014; Conroy et al., 2015). Recent results from the BEST in CLASS efficacy trial indicated moderate effect sizes for a variety of child outcomes (Sutherland et al., 2016) and moderate to large effect sizes for teacher outcomes (Conroy et al., 2016). Finally, in the only study to date examining treatment integrity of the BEST in CLASS intervention, Sutherland, Conroy, Vo, and

Ladwig (2015) found that coaches implemented the intervention with integrity, and teachers in the BEST in CLASS condition significantly increased their adherence at both post-treatment and maintenance, while BEST in CLASS teachers also had significantly higher competence ratings at post-treatment than did teachers in the comparison condition.

Rationale for the Current Study

Given the challenges teachers face in teaching young children who exhibit chronic problem behavior (Hemmeter et al., 2006) and the complexities inherent in delivering prevention programming in early childhood classrooms (Durlak, 2010), we focused our study on classroom and teacher factors hypothesized from the literature to influence treatment integrity in BEST in CLASS. Domitrovich et al.'s (2008) conceptual framework suggests that classroom climate, represented by social, psychological and/or educational aspects of the classroom environment, as well as levels of student misconduct may impact EBP delivery. Supporting this assertion, Wanless, Rimm-Kaufman, Abry, Larsen, and Patton (2015) found that teachers with higher emotional support in the classroom were more engaged during Responsive Classroom® training, and had higher observed implementation of Responsive Classroom® practices two years later. In a similar vein, Pas, Waasdorp, and Bradshaw (2015) found that levels of student problem behavior negatively impacted teacher adoption of behavioral strategies. In the current study we thus hypothesized that initial ratings of lower classroom organization, emotional support and instructional support (i.e., classroom climate) and higher initial levels of child problem behavior would predict lower adherence and competence of delivery of BEST in CLASS.

At the teacher level we focused on teacher self-efficacy and educational background as factors potentially associated with teacher delivery of BEST in CLASS. Teacher self-efficacy is defined as a teacher's judgment about their ability to promote student learning (Guo et al., 2012), and Pas, Bradshaw, and Hershfeldt (2012) point out that a teacher's belief they can successfully teach children who exhibit risk factors (e.g., behavioral, environmental) is an important component of teacher self-efficacy. Han and Weiss (2005) noted that teacher self-efficacy is related to teachers' integrity of new EBPs in their classrooms, and research is beginning to examine the relation between teacher self-efficacy and treatment integrity of prevention programs, although findings are mixed. For example, Little et al. (2013) found that teacher self-efficacy partially mediated the effect of training on treatment integrity in a dissemination trial of a school-based prevention program. However, recently Williford, Wolcott, Whittaker, and Locasale-Crouch (2015) did not find an association between teacher self-efficacy and teacher integrity of Banking Time, a tier-2 early childhood intervention. In the current study, we hypothesized that teachers who had higher levels of self-efficacy at the outset of the study would be more likely to implement BEST in CLASS extensively and with competence than would teachers who had lower levels of self-efficacy.

Teachers' education background has also been hypothesized to be associated with program delivery (Domitrovich et al., 2008; Durlak, 2010) and research suggests that teachers' education is related to program quality (Pianta et al., 2005). Williford et al. (2015) found that teachers with an early childhood education major provided greater dosage of Banking

Time than teachers without an early childhood degree. Therefore, in the current study we hypothesized that teachers with more education would implement BEST in CLASS more extensively and with greater competence than would teachers with less training. Finally, we assessed whether treatment integrity changed over time using growth curve modeling. We hypothesized that treatment adherence and competence would increase over time because the teachers were receiving practice-based coaching on practices weekly (Sutherland et al., 2015). We thus sought to investigate whether treatment integrity changed over time and whether patterns of change differed across adherence and competence.

Method

Design

This study took place in federally- or state-funded early childhood classrooms (96.0%) and in locally or privately funded programs (4.0%) in two southeastern US states. Data for the current study were drawn from intervention classrooms that were part of a larger randomized controlled trial study of BEST in CLASS. The federal or state-funded early childhood classrooms (e.g. Head Start, state funded PK) served income eligible families and children who were at risk for school failure. Classrooms were located in urban, suburban, and rural communities as either part of a local district elementary school or an early childhood education center ($n = 78$). Early childhood classrooms had on average 2.20 adults per classroom ($SD = .45$) and the mean number of children was 17.45 ($SD = 2.17$). A variety of common early childhood curricula were used in these classrooms, including the Creative Curriculum for Preschool®, the High Scope Early Childhood Curriculum, and Second Step: Social-emotional Skills for Early Learning.

Early childhood teachers working in programs serving young children (aged 3 – 5 years) were recruited to participate in the parent study. After obtaining teacher consent, teachers nominated five children in their classrooms who displayed chronic problem behavior. Once parent or guardian consent was then obtained for nominated children, screening for risk for EBD and developmental delays took place using the *Early Screening Project* (*ESP*; Feil, Severson, & Walker, 1998) Stages 1 and 2 and the *Battelle Developmental Inventory, Second Edition Screener* (*BDI II Screener*; Newborg, 2005), respectively. Children who were identified as potentially having a developmental delay were excluded from the sample; while children who screened in as at risk for EBD were included. After screening, 1–3 children per classroom, depending upon returned consents and scores on screening measures, participated in the study. Random assignment to condition (BEST in CLASS vs. comparison) occurred at the teacher level within schools at each of the two research sites.

As part of the BEST in CLASS parent study, trained observers collected adherence and competence of delivery data at eight time points over 18 weeks during the school year: baseline ($n = 1$), approximately every other week during the 14 weeks of implementation ($n = 6$), and at one-month maintenance ($n = 1$). Observation sessions lasted 10–15 minutes and were conducted on each teacher-child dyad during teacher led instructional activities (i.e., whole group or small group). Since we were interested in examining the relation between classroom factors, teacher factors, and the delivery of BEST in CLASS, adherence and competence data from the six time points that occurred during the implementation of the

intervention were included in our analyses, as well as baseline measures of classroom climate, child problem behavior, teacher self efficacy and teacher educational background.

Participants

Data from the four years of the parent study included 185 teachers (n = 92 BEST in CLASS intervention, n = 93 comparison) and 465 children (n = 231 BEST in CLASS intervention, n = 234 comparison). The current study only includes teachers and children who received the BEST in CLASS intervention.

Teachers in the study were 98.9% female (43.5% African American, 48.9% Caucasian, 3.3% Hispanic, 1.1 Asian/Pacific Islander, 2.2% other, and 1.1% unknown). Children, 62.8% of whom were male, were 50.10 months old on average (range = 36 – 60 months, SD = 6.40 months). Children were 65.4% African American, 15.6% Caucasian, 4.8% Hispanic, .4% Native American, 6.9% other, and 6.9% unknown. Teachers had an array of educational backgrounds: 1.1% high school diploma, 30.4% associate's degree, 40.2% bachelor's degree, 25.0% master's degree, 1.1% doctoral degree, and 2.2% other. Teachers averaged 11.38 years of teaching experience at the beginning of the intervention (SD = 8.98; range = 0–38.00).

Procedure

BEST in CLASS teachers were trained to use key instructional practices, including (1) rules, (2) precorrection, (3) opportunities to respond, (4) behavior specific praise, (5) corrective feedback, and (6) instructive feedback. Teachers were also trained to effectively and efficiently link the application of learned practices to identified children. Over 14-weeks, trained coaches used a practice-based coaching model to provide performance-based feedback each week to teachers on their delivery of practices, with the introduction of new practices occurring approximately every two weeks. Therefore, teachers received training on all practices in the workshop and then received coaching sequentially on each specific practice (e.g., practice-based coaching in rules for two weeks, then precorrection for two weeks, and so forth).

Measures

BEST in CLASS Adherence and Competence Scale (BiCACS).—The 14-item BiCACS (Sutherland et al., 2014) includes two subscales that assess the quantity (Adherence subscale; 7 items) and quality (Competence subscale; 7 items) of the key instructional practices (e.g., Rules, Precorrection, Opportunities to Respond, Behavior Specific Praise, Instructive Feedback, Corrective Feedback) found in the BEST in CLASS intervention (see Sutherland et al., 2014; 2015 for a detailed description of the measure). The seven items[1] on the Adherence subscale are scored on a 7-point Likert-type extensiveness scale (1 = Not at all, 3 = Some, 5 = Considerably, 7 = Extensive). When scoring adherence items coders consider extensiveness of use of each practice (see Hogue et al, 1996). The seven items on the Competence subscale are scored on a 7point Likert-type competence scale (1 = Very Poor, 3 = Acceptable, 5 = Good, 7 = Excellent). Competence is only scored on observed (i.e., adherence) items. When scoring competence, coders consider the skillfulness (e.g.,

timing, developmentally appropriate language) and responsiveness (e.g., taking into account child's individual needs) of the practice (Carroll et al., 2000).

Prior to using the BiCACS, observers participated in a 2-hour training on procedures for administering and scoring the measure and were provided a manual to facilitate scoring in the field. Following training, observers coded video recordings until reaching reliability criterion (80.0% agreement) on signal detection (i.e. agreement on whether or not observers actually coded an item) of items across three consecutive 15-minute videos consensus coded by the first and third authors. Reliability was assessed using secondary observers for 23.8% of the 1,132 total observations. Single measure intraclass correlation coefficients, ICC (2,1), were .74 and .54 for the Adherence and Competence scales, respectively.

Caregiver Teacher Report Form (C-TRF).—The C-TRF (Achenbach & Rescorla, 2000) is a teacher report measure used to assess internalizing and externalizing problem behavior in children. The 100-item instrument is intended for children aged 1.5 to 5 years and has three subscales: Externalizing, Internalizing, and Total Problems. Only the Externalizing scale was used in the current study. Items on these scales are scored using a 3-point Likert scale (i.e. 0- *not true, 1- somewhat true, and 2- often true).* In the current sample, the internal consistency with Cronbach's alpha was .92 for the Externalizing scale.

Classroom Assessment Scoring System (CLASS).—The CLASS (Pianta, LaParo, & Hamre, 2008) was used to assess classroom quality across three domains: Emotional Support, Classroom Organization, and Instructional Support. The CLASS domains consist of a total of 10 dimensions of classroom quality (e.g. *positive climate, behavior management, quality of feedback*), which are rated on a scale from 1 to 7. Dimensions were rated during classroom observations ranging from 10–20 minutes across four cycles. Trained and certified observers conducted CLASS observations at pre- and post-test. Observers participated in a 2-day training workshop led by a certified CLASS trainer and completed the reliability test required for initial certification. Inter-rater agreement data were collected on 20.7% of all CLASS observations using a secondary observer. The mean inter-rater agreement was 92.5%. To assess overall classroom quality, the dimension scores for each domain were averaged. The internal consistency for the current sample with Cronbach's alpha was .88 for Emotional Support, .89 for Classroom Organization, and .78 for Instructional Support.

Teachers' Sense of Efficacy Scale – Long Form (TSES).—The TSES (Tschannen-Moran & Hoy, 2001) is a 24-item self-report of teacher self-efficacy that has three subscales: Instructional Strategies, Classroom Management, and Student Engagement. Items are scored on a 9-point Likert-type scale (1= Nothing, 3 = Very Little, 5 Some Influence, 7 = Quite a Bit, 9 = A Great Deal). Internal consistency, calculated with Cronbach alpa coefficients for each of the three subscales, ranged from .88 to .90 for the current sample.

Teacher Education.—As part of a demographic survey, teacher's reported their own education levels, which were classified in one of three categories: High school diploma or Associates degree, Bachelor's degree, or Masters degree or higher.

Data Analyses

Latent growth curve modeling in M*plus* 8.0 (Muthén & Muthén, 1998–2017) was used to test study hypotheses. Latent growth curve modeling is the appropriate statistical method because the dependent variables, the BiCACS Adherence and Competence subscales, were measured across time and thus represent change over the course of the intervention (treatment integrity assessments were collected approximately every two weeks during the intervention phase, with the first observation occurring at the week 2 time point of intervention delivery). All predictors (Teacher Education, C-TRF, CLASS and TSES) were collected prior to intervention delivery. The data have a 3-level structure, with nesting of children within teachers, which are nested within schools. However, as there were a limited number of schools, we used a method in M*plus* called complex two-level analysis. The complex two-level procedure used a two-level (children within classrooms) model, maximum likelihood estimation of parameters, and robust standard errors and chi-square statistics that take into account non-normality and nesting of teachers within classrooms.

Inspection of plots of the adherence means and the competence means against time points indicated change in the means was approximately linear over the time points for both variables (i.e., adherence and competence). Therefore, we used a linear growth curve model with time points coded 0 to 5 so that the intercept latent variable represents status at time point 0 (i.e., beginning of the intervention phase). The slope latent variable represents the growth rate over the time points. There are two options that could be used to include the predictors in the estimation of the model. In the default for M*plus* the likelihood function is calculated from the dependent variables only. Using this method would result in dropping cases with missing scores on one or more of predictors (none of the variables in the study had more than 5% missing data). The alternative is to calculate the likelihood function from the predictors and the dependent variables. This results in including data from all participants in the analysis. We used the latter procedure.

For the Competence and Adherence subscales, we estimated a model in which the status latent variable was regressed on the pretest variables and the growth latent variable was regressed of the status latent variable and the pretest variables. For continuous pretest variables, both unstandardized and standardized regression coefficients are reported. For the variables categorizing teachers' degrees, unstandardized coefficients and Cohen's (1988) *d* are reported. Cohen's *d* was calculated following recommendations by Olejnik and Algina (2000). Cohen's *d* is often interpreted using guidelines in which .2 is considered small, .5 medium, and .8 large.

Results

BiCACS Adherence.

For BiCACS Adherence the fit of the model was not adequate. The chi-square test of fit was significant, $\chi^2(68) = 97.421$, $p = .011$. In addition, $TLI = .900$ and were $CEF = .946$ smaller than the commonly-used standard of .95. However, $RMSEAN = .043$ indicated adequate fit. In addition, at the child-level, the estimated variance of the growth latent variable was very close to zero and negative. The model was revised by removing the growth latent variable at

the child-level. The chi-square test of fit and *TLI* again indicated inadequate fit. Both *CFI* = .952 and *RMSEA* = .040 indicated adequate fit. To improve fit, four pairs of residuals were specified to be correlated. Model fit was adequate, $\chi^2(68) = 79.552$, $p = .160$, *TLI* = .961, and *RMSEA* = .027. The mean for the growth latent variable was .178 and was statistically significant, $z = 7.668$, $p = .000$ indicating that on average teacher adherence increased over time. For the 92 teachers, only three had estimated growth latent variable scores that were negative, suggesting that almost all teachers exhibited increasing adherence over time.

Results of the regression equations for the BICACS Adherence are presented in Table 1. The only predictor included in the child-level regression equation was Externalizing Problems because it is the only pretest variable that is measured at the child level. At the child-level, results are available only for the status latent variable because the growth latent variable was excluded for the model at the child-level. The results when the status latent variable is the dependent variable indicate that teachers with master's degrees or above had higher adherence at the beginning of the intervention than did teachers with associates or high school degrees. Cohen's *d* indicates the difference between the two groups was substantial. In the equation for the growth latent variable, the regression coefficient for Classroom Organization was statistically significant and negative indicating that teachers with higher pretest Classroom Organization exhibited lower growth in adherence. The coefficient for Emotional Support was statistically significant and positive, indicating that teachers with higher Emotional Support at pretest exhibited higher growth in adherence. The standardized coefficients suggest these relationships were also substantial.

BiCACS Competence.

For competence the fit of the model was not adequate. The chi-square test of fit was significant, $\chi^2(68) = 120.774$, $p = .000$. In addition, *TLI* = .837 and *CFI* = .0912 were smaller than the contemporary standard of .95. However, *RMASE* = .058 indicated adequate fit. At the child-level, the estimated variance of the growth latent variable was very close to zero and negative. The model was revised by removing the growth latent variable at the child-level. The chi-square test of fit, *TLI*, and *CFI* again indicated inadequate fit. To improve fit, one pair of residuals was specified to be correlated at the child level and four pairs of residuals were specified to be correlated at the teacher-level. The chi-square test, $\chi^2(66) = 92.104$, $p = .019$. was significant. Both *CFI* = .956 and RMSEA = .041 indicated adequate fit. *TLI* = .917 was below, but close to, the contemporary standard of .95, and above the historical standard of .90.

The mean of the growth latent variable was .163 and was statistically significant, $z = 6.936$, $p = .000$ indicating that on average competence improved over time. For the 92 teachers, only seven had estimated growth latent variable scores that were negative, suggesting that almost all teachers exhibited increasing competence over time.

Results of the regression analysis for competence are presented in Table 2. In the results when the status latent variable is the dependent variable, the regression coefficients for the two education variables were significant and positive indicating that teachers with Associates or High School degrees had lower competence at the beginning of the intervention than did teachers with Bachelor degrees or teachers with Master's degrees or

higher. Cohen's *d* indicates the differences were moderate. The regression coefficient for the status latent variable was significant and negative, indicating that teachers exhibiting lower competence at the beginning of the intervention tended to exhibit faster growth in competence. The standardized coefficient suggests the relationship is moderate. The regression coefficient for Emotional Support was statistically significant and positive and, based on the standardized regression coefficient, moderate in size. Finally the regression coefficient for Student Engagement was statistically significant and negative and, based on the standardized regression coefficient, moderate in size.

Discussion

The purpose of this study was to investigate factors associated with teacher treatment integrity of BEST in CLASS, a tier-2 prevention program that is designed to ameliorate young children's risk for developing EBDs. Results indicated that both adherence and competence increased across six observed time points. Teachers with higher levels of education delivered the program with more adherence and competence initially; however, teachers with lower levels of initial competence exhibited more growth in competence over time than teachers with higher levels of initial competence. Teachers with higher initial scores on the Classroom Organization subscale of the CLASS exhibited lower growth in adherence across time, while teachers with higher initial scores on the Emotional Support subscale implemented the program with greater competence and had higher growth in both adherence and competence across time. Contrary to hypotheses, teacher self-efficacy did not predict adherence, and teachers who reported higher initial levels of student engagement self-efficacy exhibited lower growth in competence over time.

As teachers often struggle to deliver programs designed to prevent and ameliorate chronic problem behaviors of young children with integrity (Domitrovich et al., 2010; Durlak, 2010) identifying variables that are associated with variations in delivery over time is an important goal for the field. Findings from the current study add to the literature examining the relationship between classroom-, teacher-, and child-level factors and teacher delivery of EBPs in several important ways. First, results indicated that teachers with higher levels of education delivered BEST in CLASS with greater adherence and competence at the beginning of the study. These finding are consistent with previous research that found teachers' education was related to greater integrity of the *Banking Time* program (Williford et al., 2015), although other research has not found similar relations between teacher professional characteristics and treatment integrity (Baker et al., 2010). In the current study the fact that teachers with more education demonstrated higher adherence and competence at the beginning of the study suggests that more education may help teachers transfer information more rapidly from the initial training to their classroom; these results also emphasize the importance of practice-based coaching as a support for all teachers, but in particular those with lower levels of initial preparation. This assertion is supported by the finding that teachers with lower levels of initial competence had greater growth in competence of delivery than did teachers with higher levels of initial competence. In this respect, interventions that utilize practice based coaching, such as BEST in CLASS, may have greater impacts with teachers who have less education or lower initial quality of instructional practice delivery.

Related to classroom factors, results suggest that teachers with higher initial ratings on the Classroom Organization subscale of the CLASS had lower growth in adherence than did teachers with lower initial ratings. It is plausible that teachers with higher ratings of Classroom Organization were already performing many of the BEST in CLASS practices related to classroom organization (e.g., Rules, Praise, OTR) and thus had less need (and capacity) for growth in these areas. Interestingly, however, teachers with higher initial ratings of Emotional Support delivered the program with greater competence initially and had greater growth in both adherence and competence over time. Emotional Support is characterized by the teacher's level of sensitivity toward children as well as the positive climate in the classroom. It is possible that teachers with higher initial levels of Emotional Support may have attempted to improve the quality of delivery of BEST in CLASS practices because they anticipated that the practices would exert a positive impact on their children's social and emotional development. Emotional Support does seem to be a particularly important dimension of classroom quality relating to teacher delivery, as other researchers (Wanless et al, 2015) reported on the higher engagement during Responsive Classroom® training of teachers with higher Emotional Support.

This study extends a scant literature examining associations between teacher self-efficacy and teacher delivery of prevention programs targeting social-emotional and related outcomes in early childhood settings. Contrary to hypotheses teachers with higher levels of self-efficacy did not exhibit higher initial implementation of BEST in CLASS. In fact, teachers with higher initial levels of student engagement self-efficacy exhibited lower growth in competence. These findings may be related to the nature of adherence and competence as distinct dimensions of treatment integrity (Sutherland, McLeod, Conroy, & Cox, 2013). Pas et al. (2012) suggest that a teacher's belief that s/he can successfully teach children is an important component of teacher self-efficacy. Thus, it is possible that teachers with higher student engagement self-efficacy were already aware of how their instructional behavior was related to engaging children in learning, and had less room for growth in terms of implementing the BEST in CLASS practices. Data from the current study suggest that teacher self-efficacy may not be related to adherence but may be linked to competence. Since the relations between these two dimensions of treatment integrity with child outcomes remains largely unknown (Durlak, 2010; Wolery, 2011), more research is needed to better understand these relations in order to advance the science of treatment integrity of EBPs in early childhood settings.

Finally, our hypothesis that child problem behavior would be associated with teacher implementation of BEST in CLASS was not supported. While surprising, this finding may be a result of a lack of variability in the problem behavior of the children within this tier 2 intervention; that is, all children were screened into the study based upon their elevated rates of problem behavior and risk for EBDs, and therefore all teachers in the study were focusing their practices on children with similar levels of problem behavior. Conceptual models (e.g., Domitrovich et al., 2008) that suggest levels of child problem behavior impact teacher implementation may be more relevant for universal interventions (e.g. Pas et al., 2015) than more targeted interventions such as BEST in CLASS.

Limitations

Key limitations of the current study should be kept in mind. First, implementation models (e.g., Domitrovich et al., 2008; Han & Weiss, 2005) identify factors hypothesized to influence implementation at a variety of ecological levels. Unfortunately, the current study only examined a certain number of factors at the proximal (e.g. classroom and teacher) level; future work should include macro-level factors (e.g., policies, funding), school and program level factors (e.g., school culture and climate, administrative support and leadership), and additional teacher level factors (e.g., readiness to change, attitudes towards EBPs) in order to more accurately determine how these different factors may impact implementation integrity. Second, the ICC for the BiCACS Competence subscale was modest (.54; see Cicchetti, 1994). This estimate is consistent with previous studies (e.g., Barber, Crits-Christoph, & Luborsky, 1996; Hogue et al., 2008), while lower than other estimates (e.g., Carroll et al., 2000; McLeod et al., 2016). Moreover, inter-rater reliability tends to be lower for competence than adherence (e.g., Carroll et al., 2000; Hogue et al., 2008).

Conclusion

This study adds to the literature on the delivery of prevention programs in early childhood classrooms, and specifically those that focus upon the prevention and amelioration of chronic problem behavior that place young children at risk for EBD. While there is increasing evidence that EBPs in schools, including early childhood programs, can have positive effects on children and students' academic and behavioral functioning (Durlak et al., 2011), the field continues to struggle with sustaining high-quality delivery by teachers and other school personnel. More research is clearly needed to better understand the relations between classroom and teacher factors and treatment integrity, as well as the relations between different dimensions of integrity and child outcomes. The current study is a step in the process of better understanding these relations, one that will hopefully lead to more efficient and sustainable models of EBP delivery in early childhood settings.

Acknowledgments

Funding. This research was supported by a grant (R324A110173) from the U. S. Department of Education, Institute for Education Sciences, with additional support from an NIH/NCATS Clinical and Translational Science Award to the University of Florida UL1 TR000064. The opinions expressed by the authors are not necessarily reflective of the position of or endorsed by the U. S. Department of Education.

References

Achenbach TM, & Rescorla LA (2000). CBCL/1, 5–5 & C-TRF/1, 5–5 Profiles. Research Center for Children, Youth, and Families, Burlington: VT.

Baker CN, Kupersmidt JB, Voegler-Lee ME, Arnold D, & Willoughby MT (2010). Predicting teacher participation in a classroom-based, integrated preventive intervention for preschoolers. Early Childhood Research Quarterly, 25, 270–283. doi:10.1016/j.ecresq.2009.09.005 [PubMed: 21103189]

Barber JP, Crits-Christoph P, & Luborsky L (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. Journal of Consulting and Clinical Psychology, 64(3), 619–622. doi: 10.1037/0022-006X.64.3.619 [PubMed: 8698958]

Berlin LJ, Brooks-Gunn J, McCarton C, & McCormick MC (1998). The effectiveness of early intervention: Examining risk factors and pathways to enhanced development. Preventive Medicine, 27, 238–245. doi:10.1006/pmed.1998.0282 [PubMed: 9579002]

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

Bronfenbrenner U (2005). Making human beings human: Bioecological perspectives on human development. Thousand Oaks, CA; Sage Publications

Carroll KM, & Nuro KF (2002). One size cannot fit all: A stage model for psychotherapy manual development. Clinical Psychology: Science and Practice, 9, 396–406. doi:10.1093/clipsy.9.4.396

Carroll KM, Nich C, Sifty RL, Nuro KF, Frankforter TL, Ball SA, … & Rounsaville BJ (2000). A general system for evaluating therapist adherence and competence in psychotherapy research. Drug and Alcohol Dependence, 57, 225–238. doi:10.1016/S0376-8716(99)00049-6 [PubMed: 10661673]

Carter AS, Wagmiller RJ, Gray SA, McCarthy KJ, Horwitz SM, & Briggs-Gowan MJ (2010). Prevalence of DSM-IV disorder in a representative, healthy birth cohort at school entry: sociodemographic risks and social adaptation. Journal of the American Academy of Child & Adolescent Psychiatry, 49, 686–698. doi:10.1016/j.jaac.2010.03.018 [PubMed: 20610138]

Cicchetti DV (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4), 284–290. doi: 10.1037/1040-3590.6.4.284

Conroy MA, Sutherland KS, Algina J, Ladwig C, Werch B, Martinez J, Jessee G, & Gyure M (2016). BEST in CLASS: A professional development intervention fostering high quality classroom experiences for young children with problem behavior. Manuscript submitted for publication.

Conroy MA, Sutherland KS, Algina JJ, Wilson RE, Martinez JR, & Whalon KJ (2015). Measuring teacher implementation of the BEST in CLASS intervention program and corollary child outcomes. Journal of Emotional and Behavioral Disorders, 23(3), 144–155.

Conroy MA, Sutherland KS, Vo AK, Carr SE, & Ogston P (2014). Early childhood teachers' use of effective instructional practices and the collateral effects on young children's behavior. Journal of Positive Behavior Interventions, 16, 81–92. doi: 10.1177/1098300713478666

Dane AV, & Schneider BH (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control. Clinical Psychology Review, 18, 23–45. doi:10.1016/S0272-7358(97)00043-3 [PubMed: 9455622]

Domitrovich CE, Bradshaw CP, Poduska JM, Hoagwood K, Buckley JA, Olin S, Hunter Romanelli L, Leaf PJ, Greenberg MT, & Ialongo NS (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. [PubMed: 27182282]

Domitrovich CE, Cortes RC, & Greenberg MT (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. The Journal of Primary Prevention, 28, 67–91. doi:10.1007/s10935-007-0081-0 [PubMed: 17265130]

Domitrovich CE, Gest SD, Jones D, Gill S, & DeRousie RMS (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. Early Childhood Research Quarterly, 25, 284–298. doi:10.1016/j.ecresq.2010.04.001 [PubMed: 22844183]

Domitrovich CE, Pas ET, Bradshaw CP, Becker KD, Keperling JP, Embry DD, & Ialongo N (2015). Individual and school organizational factors that influence implementation of the PAX Good Behavior Game Intervention. Prevention Science, 16(8), 1064–1074. doi:10.1007/s11121-015-0557-8 [PubMed: 25946968]

Durlak JA (2015). Studying program implementation is not easy but it is essential. Prevention Science, 16(8), 1123–1127. [PubMed: 26399607]

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

Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, & Schellinger KB (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. Child Development, 82, 405–432. doi:10.1111/j.1467-8624.2010.01564.x [PubMed: 21291449]

Feil EG, Severson HH, & Walker HM (1998). Screening for Emotional and Behavioral Delays: The Early Screening Project. Journal of Early Intervention, 21, 252–266. doi: 10.1177/105381519802100306

Fixsen DL, Naoom SF, Blasé KA, Friedman RM, & Wallace F (2005). Implementation research: A synthesis of the literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Guo Y, Connor CM, Yang Y, Roehrig AD, & Morrison FJ (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. The Elementary School Journal, 113, 3–24. doi:10.1086/665816

Han SS, & Weiss B (2005). Sustainability of teacher implementation of school-based mental health programs. Journal of Abnormal Child Psychology, 33, 665–679. doi:10.1007/s10802-005-7646-2 [PubMed: 16328743]

Hemmeter ML, Corso R, & Cheatham G (2006). A national survey of early childhood educators: Training needs and strategies. Paper presented at the Conference on Research Innovations in Early Intervention Dan Diego, CA.

Hogue A, Liddle HA, & Rowe C (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. Psychotherapy: Theory, Research, Practice, Training, 33, 332–345. doi:10.1037/0033-3204.33.2.332

Little MA, Sussman S, Sun P, & Rohrbach LA (2013). The effects of implementation fidelity in the Towards No Drug Abuse dissemination trial. Health Education, 113, 281–296. doi: 10.1108/09654281311329231

McCabe PC, & Altamura M (2011). Empirically valid strategies to improve social and emotional competence of preschool children. Psychology in the Schools, 48, 513–540. doi:10.1002/pits. 20570

McLeod BD, Southam-Gerow MA, Rodriguez A, Quinoy A, Arnold C, Kendall PC, & Weisz J,R (2016). Development and initial psychometrics for a therapist competence instrument for CBT for youth with anxiety Journal of Clinical and Adolescent Psychology Advance online publication. doi; 10.1080/15374416.2016.1253018

McLeod BD, Southam-Gerow MA, Tully CB, Rodriguez A, & Smith MM (2013). Making a case for treatment integrity as a psychological treatment quality indicator. Clinical Psychology: Science and Practice, 20, 14–32. doi:10.1111/cpsp.12020 [PubMed: 23935254]

Mendel P, Meredith LS, Schoenbaum M, Sherbourne CD, & Wells KB (2008). Interventions in organizational and community context: A framework for building evidence on dissemination and implementation in health services research. Administration and Policy in Mental Health and Mental Health Services Research, 35(1–2), 21–37. doi:10.1007/s10488-007-0144-9 [PubMed: 17990095]

Muthén LK and Muthén BO (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Nelson JR, Stage S, Duppong-Hurley K, Synhorst L, & Epstein MH (2007). Risk factors predictive of the problem behavior of children at risk for emotional and behavioral disorders. Exceptional Children, 73, 367–379. doi:10.1177/001440290707300306

Newborg J (2005). Battelle Developmental Inventory, 2nd Edition, Examiner's Manual. Rolling Meadows, IL: Riverside Publishing.

Olejnik S, & Algina J (2000). Measures of Effect Size for Comparative Studies: Applications, Interpretations, and Limitations. Contemporary Educational Psychology, 25, 241–286 [PubMed: 10873373]

Pas ET, Bradshaw CP, & Hershfeldt PA (2012). Teacher-and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support. Journal of school Psychology, 50, 129–145. doi:10.1016/j.jsp.2011.07.003 [PubMed: 22386082]

Pas ET, Waasdorp TE, & Bradshaw CP (2015). Examining contextual influences on classroom-based implementation of Positive Behavior Support strategies: Findings from a randomized controlled effectiveness trial. Prevention Science, 16(8), 1096–1106. doi:10.1007/s11121-014-0492-0 [PubMed: 24793222]

Pianta R, Howes C, Burchinal M, Bryant D, Clifford R, Early D, & Barbarin O (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality

and child-teacher interactions?. Applied developmental science, 9(3), 144–159. doi:10.1207/s1532480xads0903_2

Pianta RC, Karen M, Paro L, & Hamre BK (2008). Classroom assessment scoring system (CLASS) manual, pre-K. Baltimore, MD: Paul H. Brookes Publishing Company.

Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, Griffey R, & Hensley M (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Administration and Policy in Mental Health, 38, 65–76. [PubMed: 20957426]

Sameroff AJ (1995). General systems theories and developmental psychopathology. In Eds D. Cicchetti, D. J. Cohen, Developmental Psychopathology, Vol. 1: Theory and Methods. Oxford, England: John Wiley & Sons.

Sanetti LMH, & Kratochwill TR (2009). Toward developing a science of treatment integrity: Introduction to the special series. School Psychology Review, 38, 445–459.

Schoenwald SK, Garland AF, Chapman JE, Frazier SL, Sheidow AJ, & SouthamGerow MA (2011). Toward the effective and efficient measurement of implementation fidelity. Administration and Policy in Mental Health and Mental Health Services Research, 38, 32–43. doi:10.1007/s10488-010-0321-0 [PubMed: 20957425]

Sutherland KS, Conroy MA, Algina J, Ladwig C, Jessee G, & Gyure M (2016). Reducing child problem behaviors and improving teacher-child interactions and relationships: A randomized controlled trial of BEST in CLASS. Manuscript submitted for publication.

Sutherland KS, Conroy MA, Vo A, & Ladwig C (2015). Implementation integrity of practice-based coaching: Preliminary results from the BEST in CLASS Efficacy Trial. School Mental Health, 7, 21–33. doi:10.1007/s12310-014-9134-8

Sutherland KS, McLeod BD, Conroy MA, Abrams L, & Smith MM (2014). Preliminary psychometric properties of the BEST in CLASS Adherence and Competence Scale. Journal of Emotional and Behavioral Disorders, 22, 249–259. doi:10.1177/1063426613493669

Sutherland KS, McLeod BD, Conroy MA, & Cox JR (2013). Measuring implementation of evidence-based programs targeting young children at risk for emotional/behavioral disorders conceptual issues and recommendations. Journal of Early Intervention, 35, 129–149. doi:10.1177/1053815113515025

Tschannen-Moran M, & Hoy AW (2001). Teacher efficacy: Capturing an elusive construct. Teaching and teacher education, 17, 783–805. doi:10.1016/S0742-051X(01)00036-1

Wanless SB, Rimm-Kaufman SE, Abry T, Larsen RA, & Patton CL (2015). Engagement in training as a mechanism to understanding fidelity of implementation of the Responsive Classroom Approach. Prevention Science, 16(8), 1107–1116. [PubMed: 25351865]

Webster-Stratton C, Reid MJ, & Hammond M (2004). Treating children with early-conduct problems: Intervention outcomes for parent, child, and teacher training. Journal of Clinical Child and Adolescent Psychology, 33, 105–124. doi:10.1207/S15374424JCCP3301_11 [PubMed: 15028546]

Williford AP, Wolcott CS, Whittaker JV, & Locasale-Crouch J (2015). Program and teacher characteristics predicting the implementation of banking time with preschoolers who display disruptive behaviors. Prevention Science, 16(8), 1054–1063. doi:10.1007/s11121-015-0544-0 [PubMed: 25627344]

Wolery M (2011). Intervention research: The importance of fidelity measurement. Topics in Early Childhood Special Education, 31, 155–157. doi:10.1177/0271121411408621

Table 1

Summary of Structural Regression Analysis for Adherence

Scale	Predictor	Estimate	Standard Error	*z*	*p*	Standardized Estimate
	Dependent Variable: Status Latent Variable					
Child						
CTRF						
	Externalizing Problems	−.005	.003	−1.423	.155	−.225
Teacher						
CTRF						
	Externalizing Problems	−.034	.019	−1.800	.072	−.285
CLASS						
	Classroom Organization	−.001	.144	−0.007	.995	−.001
	Emotional Support	−.011	.182	−0.058	.954	−.015
	Instructional Support	.112	.108	1.033	.302	.151
TSES						
	Classroom Management	.119	.114	1.043	.297	.204
	Instructional Strategies	−.239	.171	−1.392	.164	−.404
	Student Engagement	.062	.159	.390	.697	.107
Degree						
	Bachelors vs Associates or High School	.225	.178	1.262	.207	.439*
	Masters or above vs Associates or High School	.639	.197	3.247	.001	1.246*
	Dependent Variable: Growth Latent Variable					
Teacher	Status Latent Variable	−.083	.045	−1.848	.065	−.332
CTRF						
	Externalizing Problems	.007	.005	1.316	.188	.220
CLASS						
	Classroom Organization	−.099	.041	−2.045	.016	−.603
	Emotional Support	.125	.054	2.321	.020	.680
	Instructional Support	−.011	.030	−0.379	.704	−.061
TSES						
	Classroom Management	.020	.034	0.576	.565	.133
	Instructional Strategies	.001	.047	0.015	.988	.005
	Student Engagement	.008	.043	0.196	.844	.058
Education						
	Bachelors vs Associates or High School	.029	.044	0.653	.514	.205*
	Masters or above vs Associates or High School	.035	.062	0.555	.579	.247*

*Cohen's *d*.

Table 2

Summary of Structural Regression Analysis for Competence

Scale	Predictor	Estimate	Standard Error	*z*	*p*	Standardized Estimate
	Dependent Variable: Status Latent Variable					
Child						
CTRF						
	Externalizing Problems	.001	.003	0.314	.753	.081
Teacher						
CTRF						
	Externalizing Problems	−.027	.022	−1.229	.219	−.211
CLASS						
	Classroom Organization	−.011	.152	−0.072	.943	−.016
	Emotional Support	.354	.174	2.032	.042	.450
	Instructional Support	.054	.130	0.418	.676	.068
TSES						
	Classroom Management	−.087	.143	−0.610	.542	−.139
	Instructional Strategies	.012	.176	0.068	.946	.019
	Student Engagement	.133	.157	0.847	.397	.213
Degree						
	Bachelors vs Associates or High School	.427	.169	2.535	.011	.776*
	Masters or above vs Associates or High School	.423	.197	2.148	.032	.768
	Dependent Variable: Growth Latent Variable					
Teacher	Status Latent Variable	−.189	.049	−3.824	.000	−.728
CTRF						
	Externalizing Problems	−.004	.006	−0.697	.486	−.116
CLASS						
	Classroom Organization	−.022	.034	−0.643	.520	−.119
	Emotional Support	.075	.034	2.192	.028	.368
	Instructional Support	−.004	.024	−0.157	.875	−.018
TSES						
	Classroom Management	.043	.023	1.846	.065	.261
	Instructional Strategies	.033	.032	1.038	.299	.199
	Student Engagement	−.067	.033	−2.017	.044	−.413
Education						
	Bachelors vs Associates or High School	.038	.036	1.054	.292	.291*
	Masters or above vs Associates or High School	.025	.046	0.537	.591	.192*

*Cohen's *d*.