Effects of CW-FIT on Teachers' Ratings of Elementary School Students at Risk for Emotional and Behavioral Disorders

Paul Caldarella, Ph.D., paul_caldarella@byu.edu

Ross A. A. Larsen, Ph.D., Ross.Larsen@byu.edu

Leslie Williams, Ed.S., leslie_williams@byu.edu

David O. McKay School of Education

Brigham Young University

Provo, UT 84602

Howard P. Wills, Ph.D., hpwills@ku.edu

Debra M. Kamps, Ph.D., dkamps@ku.edu

Juniper Gardens Children's Project

University of Kansas

444 Minnesota Ave, Suite 300

Kansas City, KS 66101

Joseph H. Wehby, Ph.D., joseph.wehby@vanderbilt.edu

Department of Special Education

Peabody College

Vanderbilt University

Nashville, TN 37203

Author Notes

Address all correspondence to Paul Caldarella, PhD, Center for the Improvement of Teacher Education and Schooling, Brigham Young University, 149D MCKB, Provo, UT 84602.

Telephone: 801-422-5081. Fax: 801-422-0199. Email: paul_caldarella@byu.edu

Acknowledgement

The research reported in this article was supported in part by a grant from the Institute of Education Sciences and the U.S. Department of Education (R324A120344), awarded to the University of Kansas, in cooperation with Brigham Young University and Vanderbilt University. The opinions presented in this article are those of the authors, and no endorsement by the agency is intended or implied.

Article citation:

Caldarella, P., Larsen, R. A., Williams, L., Wills, H. P., Kamps, D., & Wehby, J. (2017). Effects of CW-FIT on teachers' ratings of elementary school students at risk for emotional and behavioral disorders. *Journal of Positive Behavior Interventions.* Advance online publication. DOI: 10.1177/1098300717723353

Publishers' version available through:

http://journals.sagepub.com/doi/pdf/10.1177/1098300717723353

Abstract

Students with deficits in social skills have been found to experience both short- and long-term problems, including interpersonal conflicts and academic difficulties. These problems are compounded for students with emotional and behavioral disorders (EBD). Class-wide function-related intervention teams (CW-FIT), a multi-tiered classroom management program, has been shown to be effective in increasing on-task behavior and decreasing disruptive behavior of students at risk for EBD. The present study examined the effects of CW-FIT on teachers' ratings of students' social skills, problem behaviors, and academic competence. A randomized control trial was completed with 160 elementary school teachers located in 19 schools across three states. Teachers completed rating scales on 350 students identified as at risk for EBD, for whom consent had been obtained. After being randomly assigned to treatment or control conditions, CW-FIT was implemented for approximately four months in treatment classrooms, after which teachers completed post-test ratings on all students. CW-FIT implementation was associated with significantly improved teacher ratings of social skills and academic competence for students at risk for EBD, but no significant changes in teacher ratings of student problem behaviors were found. Higher fidelity of CW-FIT implementation was associated with improved outcomes. Implications, limitations, and areas for future research are addressed.

Keywords: elementary schools, emotional disturbances, behavior disorders, prevention, positive behavior intervention and supports

Effects of CW-FIT on Teachers' Ratings of Elementary School Students at Risk for Emotional and Behavioral Disorders

The importance of improving school-based treatment for students with or at risk for emotional and behavioral disorders (EBD) has been widely noted (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005). Although 10-20% of school-age children experience significant emotional and behavioral problems (Mash & Dozois, 2002), less than 1% have been identified with EBD (Wagner et al., 2005). Unfortunately, such students are often not identified until they have exhibited serious problems over multiple school years (Kauffman, 2005; Reid, Gonzalez, Nordness, Trout, & Epstein, 2004; Wagner et al., 2006). Early evidence-based interventions are needed to prevent students who are at risk for EBD from experiencing the difficulties typical of students with this disorder (Conroy, Hendrickson, & Hester, 2004; Wagner et al., 2005).

Social Skill Deficits, Problem Behaviors, and Academic Difficulties

Students classified with EBD consistently demonstrate lower social skills than typical peers or their peers with other disabilities (Lane, Carter, Pierson, & Glaser, 2006; Wagner et al., 2005). Social skills are learned, situation-specific behaviors that result in positive interpersonal interactions (Merrell & Gimpel, 1998; Young, Caldarella, Richardson, & Young, 2012). Shores and Wehby (1999) noted that the social interactions between teachers and students with EBD often consist of more negative than positive or neutral interactions. Indeed, one of the most prevalent characteristics of EBD is an inability to build or maintain satisfactory interpersonal relationships with others (Gage, 2013; Individuals with Disabilities Education Improvement Act, 2004). This is concerning since students need to learn and use appropriate social skills to be successful in school with teachers, staff, and peers. Hemmeter, Ostrosky, and Fox (2006) noted

that social skills are a critical foundation upon which school success is built. They reported a significant negative correlation between students' social skills and their challenging behavior: As social skills increased, problem behaviors decreased. These researchers also noted that many challenging behaviors are exhibited due to a lack of experience using appropriate social skills.

Students with EBD have higher levels of problem behaviors than students identified with learning disabilities, including significantly higher rates of school absenteeism, suspension, and expulsion (Lane et al., 2006; Wagner et al., 2005). Such students tend to misinterpret neutral cues as hostile and use poor problem-solving skills, resulting in antisocial behaviors that make it difficult to develop adaptive relationships with peers and teachers (Gresham, 2002; Lane, Wehby, & Barton-Arwood, 2005). School-based programs designed to improve social interactions and teach specific social skills to reduce such problem behaviors have been shown to be effective (Cook et al., 2008).

Students with EBD also struggle academically, earning lower grades, failing more classes, repeating a grade, and dropping out of school more often than other students (Wagner et al., 2005). In a meta-analysis on the academic status of students with EBD, Reid and colleagues (2004) found that such students demonstrate significantly lower academic skills across subjects when compared to a norm group without disabilities. Further, students with EBD are often viewed as less academically competent than students with learning disabilities (Lane et al., 2006). Those with EBD may struggle due to academic skill deficits, lack of content knowledge, and limited task completion (Lane et al., 2005).

Positive Behavior Interventions and Support

Teachers can incorporate effective behavior management strategies for students with, or at-risk for, EBD by using positive behavior intervention and supports (PBIS) in their classrooms. PBIS is an evidence-based, multi-tiered framework for preventing or eliminating challenging

student behaviors using proactive strategies including direct teaching and reinforcing of appropriate social skills (Sugai et al., 2000). The guiding principles of PBIS recommend application in natural school settings, support for all students, prevention of behavioral problems, consistent improvement based on collected data, and systems level change (Carr et al., 2002; Sugai et al., 2000). PBIS core components include (a) fostering students' positive relationships and interactions with peers and adults, (b) defining and teaching clear behavioral expectations during instruction time, and (c) providing students feedback on their use of appropriate social skills throughout the school day (Benedict, Horner, & Squires, 2007; Reinke, Herman, & Stormont, 2013; Sugai & Horner, 2009).

Teachers and school staff can provide effective instruction in both social skills and academics using evidence-based behavior management techniques consistently and skillfully (Kauffman, 1999). Shores et al. (1993) called for programs that emphasize proactive practices to create positive interactions involving students with EBD, their peers, and their teachers. To support the success of such students, Benner, Kutash, Nelson, and Fisher (2013) recommended using a variety of PBIS strategies including explicit instruction and interdependent group contingencies with clear classroom expectations. These effective classroom management practices (i.e., classroom rules, increased teacher praise, student feedback, and positive reinforcement) provide a foundation for improving classroom engagement and student behavior (Conroy, Sutherland, Haydon, Stormont, & Harmon, 2008). Classroom management strategies that employ social skills instruction, group contingencies, classroom organization, rules, and other similar PBIS strategies need to be further investigated to discover the effects on the social skills of students with or at risk for EBD.

Class-Wide Function-Related Intervention Teams

Class-wide function-related intervention teams (CW-FIT; Wills et al., 2010) is a multi-tiered classroom management program to aid teachers in consistent and skillful implementation of evidence-based strategies. Specifically, it incorporates a variety of PBIS strategies implemented at tier 1 (social skills instruction, interdependent group contingencies, and praise) and tier 2 (self-management and help cards) to improve classroom management and student behavior during typical academic instruction. A number of research studies have been completed on the CW-FIT program in elementary schools demonstrating positive results. The program has been shown to be effective in increasing on-task behavior at a class level (Caldarella, Williams, Hansen, & Wills, 2015; Kamps et al., 2011; Kamps, Wills et al., 2015; Wills et al., 2010) as well as decreasing disruptive behavior of students at risk for EBD (Kamps, Conklin, & Wills, 2015; Wills, Iwaszuk, Kamps, & Shumate, 2014; Wills, Kamps, Fleming, & Hansen, 2016) when implemented in general education classrooms. These studies have also reported improvements in teacher praise rates and decreases in teacher reprimands. Favorable teacher and student ratings of social validity have also been reported. However, there has yet to be a study demonstrating the effects of the program on teachers' ratings of students' social skills. Such a study is worthwhile, given the importance of social skills in student success (Merrell & Gimpel, 1998). In addition, these past studies have focused on direct observations of discrete behaviors displayed during CW-FIT implementation, rather than teacher ratings reflecting student behaviors exhibited throughout the day. Adding such ratings can help provide a more comprehensive assessment of the effects of interventions (Whitcomb & Merrell, 2013).

Summary and Research Purpose

Students who lack appropriate social skills have been found to experience both short- and long-term problems, including interpersonal conflicts and academic difficulties. These problems are compounded for students with EBD. Students with or at risk for EBD need to learn and use

appropriate social skills if they are to be successful in school. CW-FIT incorporates strategies including social skills instruction to promote students' appropriate classroom behavior. Although prior studies have shown CW-FIT to be effective in increasing on-task behavior and decreasing disruptive behavior of students at risk for EBD, additional research is needed to examine the effects of CW-FIT on teacher ratings of social skills, problem behaviors, and academic competence of these students. Given the results of prior studies, we had four hypotheses in the current study: CW-FIT would result in improved teacher ratings of (1) social skills, (2) problem behavior, and (3) academic competence for students identified as at-risk for EBD, and that (4) higher fidelity of implementation would result in improved outcomes.

Method

Participants and Settings

Demographic data are presented in Table 1. Teacher participants included 149 general and 11 special education teachers (n = 160; 95% female, 5% male; 83% Caucasian, 11% African-American, 3% Hispanic) from all elementary grades (pre-kindergarten through six). Most teachers held bachelor's (45%) or master's (41%) degrees, and the group had an average of 9 years teaching experience (range = 0 to 44 years). Teachers completed rating scales on 350 students identified as at risk for EBD (27% female, 73% male; 43% Caucasian, 39% African-American, 15% Hispanic). Students were distributed across grade levels: 1% pre-kindergarten, 18% kindergarten, 18% first, 16% second, 20% third, 13% fourth, 11% fifth, and 4% sixth grade. School records were used to obtain student demographic data. Participants were drawn from 19 schools, mostly Title I, located in urban settings in Missouri, Utah, and Tennessee. All teachers and students were participants in a multi-site, multi-year efficacy study of CW-FIT funded by the Institute of Education Sciences, with the exception of the pre-kindergarten participants (see Jolstead et al., 2016). Teachers and students each participated for one academic year.

Measures

Systematic Screening for Behavior Disorders (SSBD): Stage 1. The SSBD (Walker & Severson, 1992) is a nationally normed multi-stage process for identifying elementary students at risk for behavior disorders. In Stage 1, teachers nominate and rank order their students who exhibit externalizing or internalizing behaviors. Stage 1 inter-rater agreement (Spearman's rho) ranges between .82 and .94, and test–retest reliability ranges between .72 and .79.

Social Skills Improvement System-Teacher Form (SSIS). The standardized norm-referenced SSIS (Gresham & Elliott, 2008) consists of three scales: Social Skills, Problem Behaviors, and Academic Competence. There are 76 items rated on a 4-point Likert scale from *never* to *almost always* and 7 items rated on a 5-point Likert scale from *lowest 10%* to *highest 10%*. Sample items include "Follows your directions," "Disobeys rules or requests," and "In reading, how does this student compare with other students?" Internal consistencies (alphas) on the SSIS scales range from .94 to .97.

Direct observations. Direct observations of students were conducted by trained research staff using the Multi-Option Observation System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995), a computer software program that allows researchers to record frequency and duration events for later analysis. During a 15-min observation of an individual student, observers recorded frequency of disruptive behaviors and duration of engagement. Disruptive behavior was defined as deliberate verbal, physical, or motor displays of inappropriate behavior interfering with a student's participation and/or the productive classroom activity of peers. Engagement was defined as a student appropriately working on the assigned/approved activity.

MOOSES has been successfully used in other research studies involving observations of student behavior (see e.g., Kamps et al., 2011; Reinke et al., 2013).

School Social Behavior Scales-Second Edition (SSBS-2). The standardized, norm-referenced SSBS-2 (Merrell, 2002) consists of two teacher-rating scales: Social Competence and Antisocial Behavior. There are 64 items rated on a 5-point Likert scale from 1 (*never*) to 5 (*frequently*). Sample items include "Offers to help other students when needed" and "Insults peers." Internal consistencies (alphas) on the SSBS-2 range from .96 to .98.

Classroom Performance Survey – Elementary (CPS-E). The CPS-E was adapted from the Classroom Performance Survey (CPS; Robin, 1998), as explained by Caldarella and colleagues (2016). The original CPS was developed due to a need within secondary schools for a brief, reliable, behavior rating scale that was effective in identifying school functioning levels for students with attention deficit hyperactivity disorder (Robin, 1998). The original CPS consisted of 20 questions on a 5-point Likert scale pertaining to a student's academic and social strengths and weaknesses. Brady, Evans, Berlin, Bunford, and Kern (2012) conducted a psychometric analysis of the CPS across 23 high schools, 875 students, and 146 teachers. They concluded that (a) the CPS was comprised of two factors, Academic Competence and Interpersonal Competence; (b) CPS scores could be interpreted similarly across genders; and (c) the CPS could serve as a progress monitoring tool for at-risk students.

We used a modified CPS elementary version, CPS-E (see Caldarella et al., 2016), consisting of 17 items across two scales: Academic Competence and Interpersonal Competence. All items are rated on a 5-point Likert scale from 1 (*always*) to 5 (*never*). Sample items include "Completes class assignments," and "Relates positively to peers." Internal consistencies (alphas) on the CPS-E range from .79 to .92 (Caldarella et al., 2016). *Lower* scores on both the Academic Competence and Interpersonal Competence scales indicate improved behavior.

Treatment fidelity. Observers completed an 18-item fidelity form during each observation period (approximately 15 total observations per classroom, in both treatment and control classrooms). They observed implementation of CW-FIT and scored each component (e.g., social skills posted, praise paired with points, teams immediately rewarded) *yes* or *no*, if it was implemented. The observers recorded *yes* if a teacher demonstrated any of these components in their classroom practice. Before utilizing CW-FIT, treatment teachers implemented 3.26% (*SD* = 3.93) of CW-FIT components in the natural course of their teaching. During the study, treatment teachers implemented CW-FIT with 92.08% fidelity (*SD* = 6.72). In control classrooms, 2.43% (*SD* = 3.84) of components were implemented during baseline and 2.72% of CW-FIT components were observed (*SD* = 3.74) during the course of the study.

Social validity. Teachers and students in treatment classrooms indicated their level of satisfaction with CW-FIT by completing a social validity questionnaire at the end of their participating year. The teacher questionnaire included 18 items [15 items rated from 1 (*very true*) to 4 (*not true*) and 3 open-ended questions] regarding ease of implementation, acceptability of CW-FIT components (use of timer, teams, points, etc.) and perceptions of effectiveness on student engagement and classroom behavior. The student questionnaire consisted of two questions rated as *yes* or *no* ("Did you like playing the CW-FIT game?" and "Do you think other kids should get to play the CW-FIT game?") and three open-ended questions assessing what aspects of CW-FIT they liked or disliked and why.

Procedures

Teacher and student identification. Districts referred schools to researchers as potential participants. Researchers approached school principals to inquire whether their teachers could benefit from a classroom management intervention. A recruitment meeting was held enabling teachers to choose to participate. Teachers who volunteered completed necessary informed

consent procedures as required by school districts and institutional review boards at the participating universities.

Participating teachers identified the time of day with the greatest number of classroom behavioral challenges, during which all data collection took place. Teachers nominated students who exhibited externalizing or internalizing behaviors on the SSBD. Parents of the top three nominated students (externalizing or internalizing) were contacted and informed parent consent/student assent obtained. To be considered at risk and qualify for participation, students had to (1) score in the above average range for problem behaviors on the SSIS and (2) be verified as at risk via direct observations using MOOSES, as done in past studies of CW-FIT (see e.g., Caldarella et al., 2015; Kamps et al., 2011). Students whose MOOSES engagement levels were below 75% or whose disruptive behaviors were above 10 occurrences for a minimum of two out of five 15-min observation sessions were considered at risk, as similar characteristics have been found in other studies of behaviorally at risk students (see e.g., Kamps, Conklin et al., 2015; Wills et al., 2014; Wills et al., 2016).

Assignment to experimental conditions. Teachers were randomly assigned to either treatment or control classrooms, after which training occurred and CW-FIT commenced in treatment classrooms. Random assignment occurred at the teacher/class level due to the teachers' implementation of the intervention. Teachers' names were stratified into grade (K-2, 3-6) and type of classroom (general education, special education) and randomly assigned by the researchers using the randomized selection function in Excel. The decision as to when to begin CW-FIT implementation in assigned classrooms was based on the stability of students' engagement behavior. Safeguards to prevent contamination included (1) direct observations in control classrooms using the same treatment fidelity form used in treatment classrooms, (2) agreements from each teacher at the recruitment meeting that those randomly drawn as control

classes would not use the intervention during the study, and (3) delayed CW-FIT training for control classes until after completion of the study.

Treatment classrooms. Teachers received training after baseline data were collected and just before CW-FIT was implemented. Research staff conducted a two-hour training session during which the rationale and logistics of CW-FIT were explained and opportunities to practice components were provided. Teachers were given scripted lessons to introduce social skills and were provided feedback as they practiced. Trainers emphasized the value of using praise and included videos of other teachers implementing CW-FIT in their classrooms. To help embed the intervention into the classroom, teachers were instructed to use it as part of their regular academic instruction, where they taught as usual, and to supplement with CW-FIT to manage behavior. Trainers coached teachers (i.e., answering questions, providing feedback on quality of CW-FIT implementation) for one to two weeks after the training until teachers were able to independently implement with fidelity as indicated by fidelity observations. Coaching length varied based on how quickly teachers were able to implement CW-FIT independently. Intervention data were collected after training and coaching were completed.

Teachers directly taught CW-FIT social skills to students during 10-min lessons by defining the skills, facilitating discussions regarding rationales for the skills, and role-playing with students with follow-up discussions. The three main social skills were "Follow Directions the First Time," "How to Get the Teacher's Attention," and "Ignore Inappropriate Behavior." Each skill included four specific steps with visual cues posted in a visible location in the classroom. Teachers could choose additional social skills to teach as needed, including "Talk in a Quiet Voice," "Keep Hands, Feet, and Objects to Self," or "Stay in Your Seat." After all social skills were taught (one skill per day), teachers briefly reviewed the skills with students each day

before starting academic lessons in which CW-FIT was used. Throughout the academic lesson, teachers referred to the social skills.

Additionally, teachers stated daily point goals to guide the students and encourage demonstration of social skills. These goals were set at approximately 75-85% of the number of time intervals that would occur during the lesson. Teachers set a timer to sound at approximately 3-min intervals, though this was adjusted based on behavioral needs in classrooms (e.g., younger students needed shorter intervals; intervals could be increased if students were consistently on task). Teachers divided their classes into teams of two to six students, based on classroom conditions and convenience, to establish a group contingency. When the timer sounded, teachers scanned each team. If every student on a team was displaying appropriate social skills and was on task at that moment, teachers praised the team and marked a point on a displayed point chart (points earned were never taken away for misbehavior). The goal of working in teams was for students to provide positive peer influence and encourage appropriate behavior. At the conclusion of the CW-FIT session, teams that met the predetermined point goal received an immediate reward. The reward was short and simple, feasible for teachers, and reinforcing to students (e.g., stickers, positive notes, simple games, dance time). Through direct instruction, goals, points, teams, and rewards (e.g., group contingency), teachers attended to appropriate behaviors frequently and taught ways for students to gain help or attention in appropriate ways, as well as ignore inappropriate behaviors. These strategies, in combination with self-management and help cards described below, address two commonly reported functions of behavior, attention (from teacher or peers) and escape (from academic tasks; Ervin et al., 2001).

For students who needed additional supports, two options were available: self-management charts and help cards. Self-management charts, on which students self-record points when the timer sounds if they are on task displaying appropriate social skills, were used to help

students who displayed attention-seeking behaviors. Help cards, which students use to solicit academic assistance from teachers or peers, were used for students engaging in escape or avoidance behaviors due to academic struggles.

Control classrooms. Control classroom teachers taught using a "business as usual" approach, implementing their typical classroom management strategies. Examples of the typical strategies used by these teachers included praise, reprimands, behavior charts, clip charts, color coded behavioral cards, token economies, PBIS tickets, ClassDojo (see www.classdojo.com), daily behavior report cards, and class rewards. As noted earlier, fidelity observations revealed that control classroom teachers implemented 2.72% of CW-FIT strategies throughout the study.

Data collection schedule. During an approximately three week baseline phase (before CW-FIT), teachers completed rating scales on each student identified as at risk. One SSIS and SSBS-2 were administered, as these measures are designed to assess student behavior displayed over a longer period of time. Items directly related to the three main social skills taught during CW-FIT ("Follow Directions the First Time," "How to Get the Teacher's Attention," and "Ignore Inappropriate Behavior") are included on both the SSIS and SSBS-2 rating scales. Three CPS-E, one per week, were administered during baseline, as this measure is designed to be used as a behavioral progress monitoring tool (Caldarella et al., 2016). During intervention, teachers completed one CPS-E per month. Approximately four months after baseline, at the end of CW-FIT implementation, teachers again completed rating scales on participating students (though only one final CPS-E), as well as the social validity questionnaire. All students in participating treatment classrooms also completed a social validity questionnaire. This process was replicated across all locations, for all three years of the study.

Design and Analysis

A causal randomized controlled trial (RCT; Moher, Schulz, & Altman, 2001) was conducted to address the research questions. The design allowed for a comparison of the effects of a treatment group using CW-FIT versus an equivalent control group using only the teachers' prior methods of classroom management (Shadish, Cook, & Campbell, 2002). To establish baseline equivalence of groups on the pre-test measures, a Robust Means Modeling (RMM) analysis was completed. The RMM analysis is similar to ANOVA in that it allows comparisons of means, but has an advantage over traditional ANOVA because it does not require normality or equality of variances (Fan & Hancock, 2012). The RMM analysis and all subsequent analyses were conducted using Mplus 7.4 (Muthén & Muthén, 1998-2015).

Several analyses were then conducted to answer the research questions: (a) a RMM analysis to test whether CW-FIT results in improved scores on the SSBS-2 and SSIS measures for students at risk for EBD; (b) a latent growth curve model studying the progress of students across all eight time points of the CPS-E measure; and (c) a multivariate linear regression analysis, simultaneously regressing fidelity of treatment on all outcomes of interest. Different statistical techniques were used to answer different questions: The RMM answered causal questions generated from the RCT, while the fidelity analysis was exploratory and should not be considered to have the same causal impact.

The assumptions for all analyses included independence of observations, which was violated in the present study because students were nested within classrooms. This fact was handled in MPLUS 7.4 by using the TYPE=COMPLEX option for the RMM and fidelity analysis and TYPE=TWOLEVEL for the latent growth curve model with teacher ID being the cluster variable. Missing data were considered missing at random and handled by the full information maximum likelihood method (FIML) in Mplus, which computes a likelihood function using only observed variables. Enders and Bandalos (2001, p. 434) explained that

"Although…the FIML algorithm does not impute missing values, this borrowing of information from the observed portion of the data is conceptually analogous to replacing missing Y data points with the conditional expectation of Y given X." FIML has been shown to have better results than other missing data techniques such as listwise deletion (Enders, 2010).

The latent growth curve model assumed the two factor structure for the CPS-E measure as previously studied (Caldarella et al., 2016), namely Academic Competence and Interpersonal Competence. Each was measured across eight time points, and thus each process had its respective intercept and slope. The latent intercept for both processes was set at the last time point to test the differences of the measures at the end of the study. Treatment was regressed on the latent intercept and latent slope for both processes to study the effect of treatment on students' progression by the end of the study and how the rate of change was affected. The fidelity analysis consisted of multivariate linear regression of fidelity on outcomes within the treatment and control groups separately, due to significant differences between groups.

Results

Baseline Equivalence

Results for the RMM analysis of baseline measures showed that none of the SSIS and SSBS measures were significantly different at baseline ($p > .05$), except for one SSBS subscale (Defiant/Disruptive, $p = .03$) which showed the treatment group was higher (worse) by 1.86 points. Latent growth curve modeling showed the CPS-E measures were also higher (worse) for the treatment group at baseline for both Academic and Interpersonal factors. The fact that the treatment group was worse than the control group at baseline on these measures increased confidence that selection effects would not explain any improvement at the end of the study.

Robust Means Modeling

Results for the RMM are presented in Table 2. The *p* values for testing whether the treatment positively affected these measures were all one-sided, corresponding to the directional hypotheses. Cohen's *d* was also reported for all comparisons as an effect size estimate. The RMM results showed statistically significant outcomes favoring the treatment group over the control group for (a) SSBS-2 Social Competence Total, (b) SSBS-2 Peer Relations, (c) SSIS Communication, (d) SSIS Cooperation, (e) SSIS Assertion, (f) SSIS Empathy, and (g) SSIS Engagement. The Cohen's *d* for these measures ranged from 0.20 to 0.27, which are considered small effect sizes. The RMM showed that none of the problem behavior measures had *p* values below or near .05, indicating no treatment effect. The RMM showed academic behavior as measured by the SSBS-2 as having statistically significant results, with a $p = .004$ and a Cohen's *d* of 0.27 showing a small effect size. There were no significant differences on the SSIS Academic Competence scale.

Latent Growth Curve Model

Changes in the CPS-E factors over time are shown in Figure 1. It is important to note that *lower* scores are preferable on this measure. All reported *p* values are one sided, corresponding to the directional hypotheses. While there were improvements in both groups over time, the statistical analyses suggested greater improvement in the treatment group than the control group. The CPS-E Academic Competence score was significantly lower (0.20 points, 0.30 standard deviation units) for the treatment group compared to the control group ($p = .011$) at the last time point, indicating that teachers rated students in the treatment group as significantly more improved than students in the control group. The effect of treatment on the slope was also statistically significant ($b = -0.03$, $p < .001$). In the control group Academic Competence scores decreased by 0.03 points (0.04 standard deviation units) for every one-unit increase in time on average. In the treatment group Academic Competence scores decreased by 0.06 points (0.09

standard deviation units) for every one-unit increase in time: Thus CW-FIT approximately doubled the natural improvement in this measure.

There was no significant difference between the treatment and control groups on CPS-E Interpersonal Competence scores ($p > .05$) at the last time point. However, the effect of treatment on the slope was statistically significant ($b = -0.03$, $p = .006$). In the control group, Interpersonal Competence decreased by 0.03 points (0.07 standard deviation units) per one-unit increase in time on average. In the treatment group Interpersonal Competence decreased by .06 points (0.14 standard deviation units) per unit of time: Thus, the effect of CW-FIT approximately doubled the natural improvement in this measure also.

Fidelity Results

Fidelity results are reported in Table 3. It was necessary to run separate analyses for the treatment and control groups as the mean difference in fidelity was high (as earlier mentioned), thus violating normality assumptions. All reported *p* values were one-sided, corresponding to the directional hypothesis. A standardized Beta (β), as produced by Mplus, was included as an effect size. In the control group, only one relation was significant (SSBS-2 Social Competence total $b = 0.41$, $p = .038$, $\beta = 0.14$) and the non-significant results are not reported. The fidelity results reported below are for the treatment group only.

Fidelity results for the treatment group showed that effects for SSBS-2 Social Competence Total, SSBS-2 Self-Management/Compliance, and CPS-E Interpersonal Competence were all significant ($p < .05$) suggesting that an increase of fidelity affected these social skill measures in a positive way. The absolute value of the standardized βs ranged from 0.15 to 0.20, which are considered small effect sizes.

Fidelity results for the treatment group also showed that *p* values were less than .05 for SSBS-2 Antisocial/Aggressive, SSIS Hyperactivity/Inattention, and SSIS Autism Spectrum,

suggesting that higher fidelity affected all of these problem behavior measures in a positive way. The absolute value of the standardized βs ranged from 0.12 to 0.18, which are considered small effect sizes.

Finally, fidelity results for the treatment group were also positive for SSBS-2 Academic Behavior and CPS-E Academic Competence ($p < .05$), suggesting that higher fidelity affected these academic measures in a positive way. The absolute value of the standardized βs ranged from 0.14 to 0.22, which are considered small effect sizes.

Social Validity

Results indicated 88% of teachers enjoyed teaching with CW-FIT; 94% found the program easy to learn and implement; 96% found teams and points helpful in improving student behavior; 92% learned new skills to help manage student behavior; 92% would use CW-FIT skills with future classes; 90% would recommend the program to colleagues; 96% reported that students liked CW-FIT; and 100% agreed that students were more focused and engaged when CW-FIT was implemented. Results indicated 93% of students liked CW-FIT, particularly the group rewards and point system; 90% of students agreed that their peers should get to use CW-FIT because it is fun and helps improve behavior and teamwork in class.

Discussion

The purpose of this study was to investigate the effects of CW-FIT on teachers' ratings of students' social skills, problem behaviors, and academic competence, as well as the impact of treatment fidelity on outcomes. A discussion of the results regarding the four research hypotheses follows. Our first hypothesis was that CW-FIT would result in improved teacher ratings of social skills for students at risk for EBD. Results showed significant post-test improvements in social skills for students in treatment classrooms compared to those in control classrooms. These results correspond to prior studies of CW-FIT which have shown

improvements in students' classroom behavior following the intervention (Caldarella et al., 2015; Kamps et al., 2011; Kamps, Wills, et al., 2015; Wills et al., 2010; Wills et al., 2016). Results make sense as CW-FIT involves social skills being directly taught and reinforced by classroom teachers. These results are important since improvements in at-risk students' social skills increases the probability of their success in school (Hemmeter et al., 2006) and can result in improved classroom learning environments for all students (Conroy et al., 2008). It is important to note that significant differences were found for the majority of social skills measured, with the exception of skills associated with self-management and self-control. It is unclear why no significant differences were found between groups on these variables.

The second hypothesis was that CW-FIT would result in improved teacher ratings of problem behaviors for students at-risk for EBD. However, no significant differences in problem behavior ratings were found at post-test. These results are somewhat surprising, since prior studies have shown significant decreases in disruptive student behaviors during direct observations (Kamps, Conklin et al., 2015; Wills et al., 2014; Wills et al., 2016). However, in these prior studies direct observations measured different discrete behaviors than were measured in the current study. For example, we measured ratings of behaviors such as blaming others for problems, taking things belonging others, and being overly demanding of teacher attention. Prior CW-FIT studies measured behaviors such as talking to peers without permission, name-calling, and throwing materials (see e.g., Kamps, Conklin et al., 2015). Additionally, when teachers completed rating scales, their ratings likely reflected student behaviors exhibited throughout the day, not being limited to behaviors only observed while the intervention was being implemented (as has been reported in prior CW-FIT studies).

The third hypothesis was that CW-FIT would result in improved teacher ratings of academic competence for students at risk for EBD. Significant improvements in treatment

students compared to control students on both the SSBS-2 and CPS-E measures supported this hypothesis. These results are important, given the critical role of such skills in students' academic success (Hemmeter et al., 2006). A possible reason the SSIS Academic Competence score was not significant may be that it consisted of ranking students compared to other students in their classrooms, which may be less sensitive to absolute changes in individual students, whereas the SSBS-2 and CPS-E use scales that do not involve such cross student comparisons.

The final hypothesis was that higher fidelity of CW-FIT implementation would result in improved outcomes. Improved fidelity was associated with better outcomes on many of the measures in the treatment group, consistent with prior studies showing the importance of CW-FIT fidelity on student outcomes (e.g., Caldarella et al., 2015; Kamps, Wills et al., 2015). Results also suggest that elementary school teachers can implement CW-FIT with fidelity, which is also consistent with these past studies. Additionally, results indicated that these practices, as measured by the fidelity observations, were not naturally occurring at high frequencies in control classrooms.

Implications

There are several implications of these findings. CW-FIT appears to be a feasible and socially valid approach to help elementary school teachers implement early, evidence-based, PBIS interventions for students at-risk for EBD. The significant increases in teacher ratings of social skills and academic competence were impressive, given the moderate length of the intervention. This seems a wise investment, considering the negative outcomes for such students if not provided treatment (Conroy et al., 2004; Wagner et al., 2005). CW-FIT is administered universally ensuring at-risk students can remain in class during academic instruction, thus decreasing the likelihood of falling behind academically. Students have the opportunity to practice social skills with peers in class rather than with a few select individuals in a pullout

group. Since one defining characteristic of EBD is the inability to build or maintain relationships, interventions such as CW-FIT that directly teach and reinforce social skills are needed to help students improve their interactions with teachers and peers.

Limitations and Directions for Future Research

There were several limitations to this study to address in future research. Teacher rating scales were the only outcome measures examined. Behavior rating scales have been criticized for being more subjective, as they are based on raters' perceptions of behavior (Cook, Volpe, & Delport, 2014). Teachers who implemented the intervention may have been particularly looking for positive results (expectancy effects). This is somewhat suggested in the decreasing baseline trends in the CPS-E scores in the treatment group (as shown in Figure 1) indicating improvement in student ratings. However, behavioral improvements were not consistent across measures (with more favorable results seen on measures of social skills and academic competence than on problem behaviors), an argument against teachers' ratings reflecting the effects of subjective positive expectancies. Future research in this area would benefit from adding other measures of student outcomes including office discipline referrals, systematic direct observations, direct behavior ratings, and standardized academic test scores. Although this study spanned three states, it was not nationally representative. The study examined only elementary school students. The teacher rating scales were not counterbalanced to control for order effects. Finally, the study followed students for only one year. Future research would benefit from addressing these limitations by including more nationally representative samples, exploring use of CW-FIT in secondary schools, counterbalancing measures, and following students over longer periods of time to determine whether there are lasting effects.

Conclusion

In conclusion, CW-FIT appears to be a promising intervention for improving outcomes for elementary students at risk for EBD. These results are important, given the critical role of social skills and academic competencies in student success. Without intervention, such students are at increased risk for a variety of negative outcomes including school failure, dropout, and later employment difficulties (Lane et al., 2006; Wagner et al., 2005). To support the academic engagement of these students, experts recommend using a variety of PBIS strategies including classroom rules and expectations, explicit instruction, increased teacher praise, student feedback, and positive reinforcement, as well as group contingencies (Benner et al., 2013; Conroy et al., 2008). Teachers can assist in prevention of and intervention for EBD by providing effective instruction in both social skills and academics using evidence-based behavior management techniques consistently and skillfully (Kauffman, 1999). The CW-FIT program, which employs such strategies in classroom settings, appears to be a viable approach for early intervention in elementary schools.

References

Benedict, E. A., Horner, R. H., & Squires, J. K. (2007). Assessment and implementation of positive behavior support in preschools. *Topics in Early Childhood Special Education, 27*, 174–192. doi:10.1177/02711214070270030801

Benner, G. J., Kutash, K., Nelson, J. R., & Fisher, M. B. (2013). Closing the achievement gap of youth with emotional and behavioral disorders through multi-tiered systems of support. *Education and Treatment of Children, 36*(3), 15–29. doi:10.1353/etc.2013.0018

Brady, C. E., Evans, S. W., Berlin, K. S., Bunford, N., & Kern, L. (2012). Evaluating school impairment with adolescents using the classroom performance survey. *School Psychology Review, 41*, 429–446.

Caldarella, P., Larsen, R. A., Williams, L., Wehby, J., Wills, H. P., & Kamps, D. (2016). Monitoring academic and social skills in elementary school: A psychometric evaluation of the Classroom Performance Survey. *Journal of Positive Behavior Interventions.* Advance online publication. DOI: 10.1177/1098300716665081

Caldarella, P., Williams, L., Hansen, B. D., & Wills, H. (2015). Managing student behavior with Class-wide function-related intervention teams: An observational study in early elementary classrooms. *Early Childhood Education Journal, 43*, 357–365. doi:10.1007/s10643-014-0664-3

Carr, E., Dunlap, G., Horner, R., Koegel, R., Turnbull, A., Sailor, W., … Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions, 4*, 4–16.

Conroy, M. A., Hendrickson, J. M., & Hester, P. P. (2004). Early identification and prevention of emotional and behavioral disorders. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur

(Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 199-215). New York, NY: Guilford.

Conroy, M. A., Sutherland, K., Haydon, T., Stormont, M., & Harmon, J. (2008). Preventing and remediating young children's chronic problem behaviors: An ecological classroom-based approach. *School Psychology Review, 46*, 3-17.

Cook, C. R., Gresham, F. M., Kern, L., Barreras, R. B., Thornton, S., & Crews, S. D. (2008). Social skills training for secondary students with emotional and/or behavioral disorders: A review and analysis of the meta-analytic literature. *Journal of Emotional and Behavioral Disorders, 16*, 131–144. doi:10.1177/1063426608314541

Cook, C. R., Volpe, R. J., & Delport, J. (2014). Systematic progress monitoring of students with emotional and behavioral disorders: The promise of change-sensitive brief behavior rating scales. In H. M. Walker & F. M. Gresham (Eds.), *Handbook of Evidence-Based Practices for Emotional and Behavioral Disorders: Applications in Schools* (pp. 211–228). New York, NY: The Guilford Press.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

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

Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, A. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review, 30*(2), 193–210.

Fan, W., & Hancock, G. R. (2012). Robust means modeling: An alternative for hypothesis testing of independent means under variance heterogeneity and nonnormality. *Journal of Educational and Behavioral Statistics, 37*(1), 137-156. doi: 10.3102/1076998610396897

Gage, N. A. (2013). Characteristics of students with emotional disturbance manifesting internalizing behaviors: A latent class analysis. *Education and Treatment of Children, 36*(4), 127–145. doi:10.1353/etc.2013.0038

Gresham, F. M. (2002). Social skills assessment and instruction for students with emotional and behavioral disorders. In K. L. Lane, F. M. Gresham, & T. E. O'Shaughnessy (Eds.), *Interventions for children with or at risk for emotional and behavioral disorders* (pp. 242–258). Boston, MA: Allyn & Bacon

Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System: Rating Scales*. Bloomington, MN: Pearson Assessments.

Hemmeter, M. L., Ostrosky, M., & Fox, L. (2006). Social and emotional foundations for early learning: A conceptual model for intervention. *School Psychology Review, 35*(4), 583–601.

Individuals with Disabilities Education Improvement Act 2004, 20 U.S.C. § 1462(h) et seq. (2004).

Jolstead, K. A., Caldarella, P., Hansen, B. D., Korth, B. B., Williams, L. & Kamps, D. (2016). Implementing positive behavior support in preschools: An exploratory study of CW-FIT tier 1. *Journal of Positive Behavior Interventions.* Advance online publication. doi: 10.1177/1098300716653226

Kamps, D., Conklin, C., & Wills, H. (2015). Use of self-management with the CW-FIT group contingency program. *Education and Treatment of Children, 38*, 1–32. doi:10.1353/etc.2015.0003

Kamps, D., Wills, H. P., Dawson-Bannister, H., Heitzman-Powell, L., Kottwitz, E., Hansen, B., & Fleming, K. (2015). Class-Wide Function-Related Intervention Teams "CW-FIT" efficacy trial outcomes. *Journal of Positive Behavior Interventions, 17,* 134–145. doi:10.1177/1098300714565244

Kamps, D., Wills, H. P., Heitzman-Powell, L., Laylin, J., Szoke, C., Petrillo, T., & Culey, A. (2011). Class-wide function-related intervention teams: Effects of group contingency programs in urban classrooms. *Journal of Positive Behavior Interventions, 13,* 154–167. doi:10.1177/1098300711398935

Kauffman, J. M. (1999). How we prevent the prevention of emotional and behavioral disorders. *Exceptional Children, 65*(4), 448-468.

Kauffman, J. M. (2005). *Characteristics of emotional and behavioral disorders of children and youth* (8th ed.). Upper Saddle River, NJ: Pearson Education.

Lane, K. L., Carter, E. W., Pierson, M. R., & Glaeser, B. C. (2006). Academic, social, and behavioral characteristics of high school students with emotional disturbances or learning disabilities. *Journal of Emotional and Behavioral Disorders, 14,* 108–117. doi:10.1177/10634266060140020101

Lane, K. L., Wehby, J., & Barton-Arwood, S. M. (2005). Students with and at risk for emotional and behavioral disorders: Meeting their social and academic needs. *Preventing Social Failure, 49*(2), 6–9. doi:10.3200/psfl.49.2.6-9

Mash, E. J., & Dozois, D. J. A. (2002). Child psychopathology: A developmental systems perspective. In E. J. Mash & R. A. Barkley (Eds.), *Child psychology* (2nd ed., pp. 3-71). New York, NY: Guilford.

Merrell, K. W. (2002). *School Social Behavior Scales-Second Edition*. Baltimore, MD: Brookes Publishing.

Merrell, K. W., & Gimpel, G. A. (1998). *Social skills of children and adolescents: Conceptualization, assessment, treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Moher, D., Schulz, K. F., & Altman, D. (2001) The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association, 285*(15), 1987-91. doi:10.1001/jama.285.15.1987

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.) Los Angeles, CA: Muthén & Muthén.

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38,* 130–143. doi:10.1177/00224669040380030101

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions, 15,* 39–50. doi:10.1177/1098300712459079

Robin, A. L. (1998). *ADHD in Adolescents.* Guilford Press: New York, NY.

Severson, H. H., Walker, H. W., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*(2), 193-223. doi:10.1016/j.jsp.2006.11.003

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton-Mifflin.

Shores, R. E., Jack, S. L., Gunter, P. L., Ellis, D. N., DeBriere, T. J., & Wehby, J. H. (1993). Classroom interactions of children with behavior disorders. *Journal of Emotional and Behavioral Disorders, 1,* 27–39. doi:10.1177/106342669300100106

Shores, R. E., & Wehby, J. H. (1999). Analyzing the classroom social behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 7,* 194–199. doi:10.1177/106342669900700401

Sugai, G., & Horner, R. H. (2009). Responsiveness-to-intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality, 17,* 223–237. doi:10.1080/09362830903235375

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., … Ruef, M. (2000). Applying positive behavior support and functional behavior assessment in schools. *Journal of Positive Behavior Interventions, 2,* 131-143. doi:10.1177/109830070000200302

Tapp, J. T., Wehby, J. H., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments, & Computers, 27,* 25–31. doi:10.3758/bf03203616

Wagner, M., Friend, M., Bursuck, W. D., Kutash, K., Duchnowski, A. J., Sumi, C., & Epstein, M. H. (2006). Educating students with emotional disturbances: A national perspective on school programs and services. *Journal of Emotional Behavioral Disorders, 14*(1), 12–30. doi:10.1177/10634266060140010201

Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13,* 79–96. doi:10.1177/10634266050130020201

Walker, H. M., & Severson, H. H. (1992). *Systematic Screening for Behavior Disorders*. Longmont, CO: Sopris West.

Whitcomb, S., & Merrell, K. W. (2013). *Behavioral, social, and emotional assessment of children and adolescents*. New York, NY: Routledge.

Wills, H. P., Iwaszuk, W. M., Kamps, D., & Shumate, E. (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children, 37,* 191–210. doi:10.1353/etc.2014.0016

Wills, H. P., Kamps, D., Hansen, B., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B. (2010). The classwide function-based intervention team program. *Preventing School Failure, 54,* 164–171. doi:10.1080/10459880903496230

Wills, H. P., Kamps, D., Fleming, K., & Hansen, B. D. (2016). Student and teacher outcomes of the class-wide function-related intervention team efficacy trial. Exceptional Children, 83(1), 58-76. doi:10.1177/0014402916658658

Young, E. L., Caldarella, P., Richardson, M. J., & Young, K. R. (2012). Positive behavior support in secondary schools: A practical guide. New York: Guilford Press.

Table 1

Descriptive Data for Teachers (n = 160), Students (n = 350), and Schools (n = 19)

Name of site	Elementary school	Number of teachers	Number of students	Percent of total sample	Free/reduced price lunch	School Size
Site 1	School 1	9	18	5.1%	71.9%	220
	School 2	4	12	3.4%	48.2%	243
	School 3	8	17	4.9%	95.8%	319
	School 4	10	20	5.7%	59.2%	578
	School 5	5	10	2.9%	81.2%	290
	School 6	9	17	4.9%	72.7%	289
	School 7	15	36	10.3%	65.0%	515
Site 2	School 8	10	32	9.1%	69.2%	425
	School 9	10	20	5.7%	81.0%	490
	School 10	5	12	3.4%	35.9%	476
	School 11	10	14	4.0%	55.3%	409
	School 12	8	18	5.1%	34.0%	630
	School 13	13	29	8.3%	82.7%	504
Site 3	School 14	8	18	5.1%	52.0%	519
	School 15	7	16	4.6%	94.2%	677
	School 16	9	23	6.6%	98.1%	475
	School 17	5	7	2.0%	40.9%	472
	School 18	8	18	5.2%	91.5%	317
	School 19	7	13	3.7%	99.0%	384

Table 2

Robust Means Modeling Results of the Effect of CW-FIT Treatment on Outcomes of Interest at Post-test (Student n=177 treatment, 134 control, 311 total)

Variable	Mean treatment post-test	Mean control post-test	*p* value of differences	Cohen's *d*
SSBS-2 Social Competence Total	99.30	91.55	.004*	0.35
Peer Relations	42.90	38.62	.002*	0.39
Self-Management/Compliance	30.99	29.59	.088	0.18
Academic Behavior	25.15	23.34	.004*	0.27
SSBS-2 Antisocial Behavior Total	78.95	77.93	.379	0.04
Hostile/Irritable	34.63	34.40	.439	0.02
Antisocial/Aggressive	22.59	21.76	.207	0.10
Defiant/Disruptive	21.68	21.64	.485	0.00
SSIS Social Skills Total	71.19	67.33	.052	0.19
Communication	11.86	11.21	.048*	0.20
Cooperation	8.47	7.62	.006*	0.27
Assertion	11.84	11.05	.008*	0.23
Responsibility	8.48	8.31	.337	0.05
Empathy	9.52	8.73	.044*	0.20
Engagement	11.97	11.07	.030*	0.24
Self-Control	9.53	9.22	.300	0.07
SSIS Problem Behaviors Total	28.46	29.68	.285	-0.09
Externalizing	13.68	13.56	.449	0.02
Bullying	3.58	3.50	.432	0.03
Hyperactivity/ Inattention	10.15	10.64	.214	-0.11
Internalizing	4.88	5.64	.100	-0.19
Autism Spectrum	15.79	16.56	.170	-0.12
SSIS Academic Competence	12.16	11.88	.365	0.04

Note. Mean scores on the SSBS-2 and SSIS are raw scores. Higher scores on the SSBS-2 and SSIS indicate higher levels of behavior. *$p < 0.05$. All *p* values are one-tailed.

Table 3

Effects of Fidelity on Outcomes of Interest in the Treatment Group (n=177)

	Effect of fidelity on response			
Response variable	*b*	*SE*	*p*	*β*
SSBS-2 Social Competence Total	0.31	0.14	.013*	0.20
Peer Relations	0.03	0.08	.340	0.03
Self-Management/ Compliance	0.14	0.08	.033*	0.15
Academic Behavior	0.45	0.26	.042*	0.14
SSBS-2 Antisocial Behavior Total	-0.26	0.20	.094	-0.15
Hostile/Irritable	-0.14	0.14	.157	-0.11
Antisocial/Aggressive	-0.19	0.11	.040*	-0.18
Defiant/Disruptive	-0.59	0.44	.087	-0.15
SSIS Social Skills Total	0.24	0.27	.187	0.08
Communication	-0.39	0.18	.987	-0.19
Cooperation	-0.01	0.11	.646	-0.01
Assertion	0.08	0.05	.052	0.16
Responsibility	0.04	0.05	.195	0.09
Empathy	0.00	0.04	.478	0.01
Engagement	0.04	0.06	.242	0.08
Self-Control	0.00	0.05	.480	0.00
SSIS Problem Behaviors Total	0.10	0.04	.985	0.18
Externalizing	0.01	0.06	.553	0.01
Bullying	-0.15	0.09	.057	-0.14
Hyperactivity/ Inattention	-0.07	0.04	.047*	-0.15
Internalizing	-0.08	0.06	.099	-0.11
Autism Spectrum	-0.06	0.04	.039*	-0.12
SSIS Academic Competence	-0.18	0.09	.974	-0.19
CPS-E Academic Competence	-0.03	0.01	.005*	-0.22
CPS-E Interpersonal Competence	-0.02	0.01	.032*	-0.15

Note. Higher scores on the SSBS-2 and the SSIS indicate higher levels of behavior, whereas lower scores on the CPS-E indicate improvement. *$p < 0.05$. All *p* values are one-tailed.

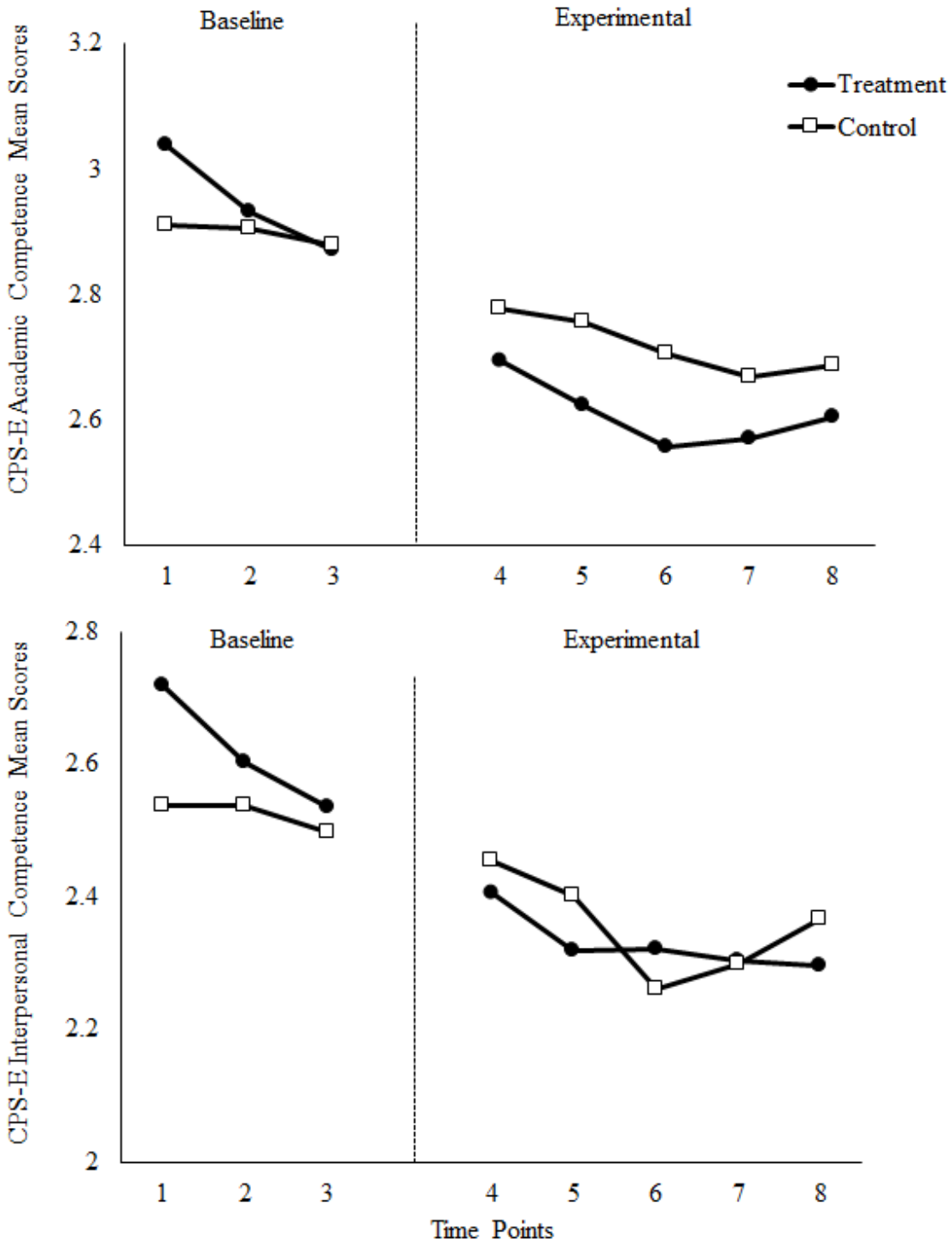

Figure 1. Mean scores of CPS-E academic and interpersonal competence factors over time. Lower scores on the CPS-E indicate improvement.