Prevention Science (2018) 19:1043–1054
https://doi.org/10.1007/s11121-018-0932-3

The Incredible Years Teacher Classroom Management Program: Outcomes from a Group Randomized Trial

Wendy M. Reinke[1] · Keith C. Herman[1] · Nianbo Dong[2]

Published online: 19 July 2018

Abstract
This group randomized controlled trial (RCT) evaluated the efficacy of the Incredible Years Teacher Classroom Management Program (IY TCM) on student social behavioral and academic outcomes among a large diverse sample of students within an urban context. Participants included 105 teachers and 1817 students in kindergarten to third grade. Three-level hierarchical linear models (HLM) were conducted to examine the overall treatment effects on teacher-reported student behavior and academic outcomes. In addition, multi-level moderation analyses were conducted to examine whether the treatment effects on student outcomes differed by demographic variables and pretest measures of social emotional and disruptive behavior and academics. Findings indicate that IY TCM reduced student emotional dysregulation ($d = -0.14$) and increased prosocial behavior ($d = 0.13$) and social competence ($d = 0.13$). In addition, students initially lower on measures of social and academic competence demonstrated significant improvements on the same measure at outcome in comparison to similar peers in control classrooms. Practical significance of the findings and implications for schools and policy makers are discussed.

Keywords Classroom management · Teacher training · Prevention · Social behavior · Academic competence

Ineffective classroom behavior management practices are associated with disruptive classroom behaviors that interfere with instruction, child development, and academic achievement. In poorly managed classrooms, students have little structure or support for consistent behavioral expectations and, as a result, students may be off task more and engage in higher rates of disruptive behaviors (Jones and Jones 2004). Negative teacher-student interactions are also more likely to occur in poorly managed classrooms (Conroy et al. 2009; Reinke and Herman 2002), and these classroom environments can contribute to students' risk for developing behavior problems (Webster-Stratton and Reid 2004). Ineffective classroom management has also been linked to long-term negative academic, behavioral, and social outcomes for students (Kellam et al. 1998; National Research Council 2002).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11121-018-0932-3) contains supplementary material, which is available to authorized users.

✉ Wendy M. Reinke
reinkew@missouri.edu

[1] Department of Educational, School, & Counseling Psychology, University of Missouri, Columbia, MO 65211, USA

[2] University of North Carolina at Chapel Hill, 110 Peabody Hall, CB 3500, Chapel Hill, NC 27599, USA

At the same time, considerable research has demonstrated that teachers can promote a positive classroom environment through the use of effective universal classroom management strategies (Curby et al. 2013; Ialongo et al. 2001; Simonsen et al. 2008). Teachers trained in effective classroom management strategies can help students who are aggressive, disruptive, and uncooperative to develop the appropriate social behavior that is a prerequisite for their success in school (Walker et al. 1995). Even though research clearly supports the use of specific classroom management practices, such as praise and proactive teaching, direct observations of teachers indicate they do not regularly use these practices (e.g. Reinke et al. 2013a). Further, many teachers find managing student disruptive behavior challenging and request additional training and support in the area of classroom management (Reinke et al. 2011). Thus, there is a need for prevention interventions that focus on training teachers to use classroom strategies that are developmentally appropriate and supportive of students' emotional and behavioral growth.

The Incredible Years Teacher Classroom Management Program

The Incredible Years Teacher Classroom Management Program (IY TCM) is a universal classroom management

program for teachers of students in pre-school through third grade. IY TCM uses social learning theory (Bandura 1977) via video modeling to increase teacher learning of new skills. During each training session, teachers view video recordings of effective strategies, role-play the use of strategies, and receive feedback from the IY TCM leaders and other teachers participating in the group. IY TCM is principle driven and therefore flexible in adapting to the skill levels and specific classroom experiences of each teacher (Webster-Stratton et al. 2011). The IY TCM incorporates teacher experiences and the cultural contexts of the participants' classroom into the workshop training content. Teachers learn key classroom management skills through discussion, observation of video-recorded examples of classroom situations, role-play rehearsal, and verbal and written assignments that are reviewed and returned (Webster-Stratton et al. 2004). In addition, the IY TCM embeds coaching within the training model. After and between each training session, teachers are followed by a coach who conducts observations, provides performance feedback, and assists with problem solving, goal setting, and implementation of strategies (Reinke et al. 2012).

Supporting Evidence for the IY TCM Program A number of randomized control trials (RCTs) of the IY series have included the TCM program. The first RCT was a prevention trial with 272 students from head start and their mothers (Webster-Stratton et al. 2001). Participants were randomly assigned to either the IY parent and teacher programs or a control condition. Results indicated that in classrooms of teachers receiving IY TCM, students were observed to have higher on-task behavior, increased pro-social behaviors, and decreased aggression. A second RCT was conducted with 159 students diagnosed with conduct problems (Webster-Stratton et al. 2004). The study compared child and parent training with and without IY TCM, resulting in five combinations of IY programs (three with IY TCM) and a no-treatment control condition. Teachers who received IY TCM were observed post-intervention to use more praise and be more nurturing, consistent, and confident than teachers who did not receive IY TCM. Additionally, students in classrooms with teachers who received IY TCM were significantly less aggressive with peers and more cooperative with teachers. A third trial of the IY TCM conducted by Webster-Stratton et al. (2008) showed similar effects in preschool settings for an IY TCM intervention combined with the IY child curriculum.

IY TCM Only Evaluations A few studies have evaluated the impact of the IY TCM program as a stand-alone intervention. Hutchings et al. (2007) using a sub-sample of students in each classroom with low and high ratings of conduct problems found that primary school teachers in Wales who received the IY TCM as compared to teachers who had not significantly increased their use of direct commands ($d = 0.94$) and praise ($d = 0.86$). Further, the students in classrooms of trained teachers were significantly more compliant ($d = 0.63$) and had more positive classroom behaviors ($d = 0.99$) than students in classrooms of untrained teachers. More recently, Hitchings, Martin-Forbes, Daley, and Williams (2013) evaluated the efficacy of the IY TCM to improve teacher behavior, student behavior classroom-wide, and with students at risk of developing conduct problems. Participants included six intervention and six control classrooms comprising 12 teachers and 107 students at risk of developing conduct problems (aged 3 to 7 years). Results showed a significant reduction in observed classroom off-task behavior ($d = 0.53$), teacher negatives to target students ($d = 0.36$), target child negatives toward the teacher ($d = 0.42$), and target child off-task behavior ($d = 0.48$). Although the results were promising, the study was limited by the small sample.

Rationale for the Current Study The IY TCM as a stand-alone intervention has great potential to positively impact large numbers of students given that each year teachers have up to 30 students in their classrooms. The prior studies discussed have shown promising results for the IY TCM program on teacher and student behavior, although none without limitations. Further, none of the IY TCM studies have investigated the potential impact of the IY TCM program on academic outcomes. By promoting student skills and reducing disruptive behaviors, IY TCM implementation may allow for more instructional time and more student time on task which could lead to more learning and higher academic performance.

The purpose of this study was to evaluate the efficacy of the IY TCM program using a large group RCT on student social emotional, disruptive behavior, and academic outcomes in schools. This is one of the first large-scale stand-alone evaluations of the IY TCM program with the full range of elementary school teachers (kindergarten to third grade) that investigates both change in student social behavior and academic outcomes for students. Based on previous research (e.g., Webster-Stratton et al. 2001; Webster-Stratton et al. 2004), we hypothesized that students in the classrooms of teachers who received the IY TCM intervention would demonstrate reductions in concentration problems, disruptive behaviors, and emotional dysregulation in comparison to students in classrooms of the control group teachers. We also expected students in classrooms of teachers who received the IY TCM intervention to demonstrate increases in prosocial behavior, overall social competence, academic competence, and academic achievement.

Lastly, there is a growing body of research suggesting that the effects of universal prevention programs may not be consistent across the full population of students exposed to the program (Farrell et al. 2013; Schochet et al. 2014). In fact, there is emerging evidence that the variation in intervention responsiveness may be predicted by baseline behavioral or

social emotional risks (e.g. Bradshaw et al. 2015; Kellam et al. 1998). Such variation is expected in a developmental epidemiological conceptual framework (Kellam et al. 1991). In this framework, the presence of a social, emotional, or behavioral problems (such as disruptive behaviors, concentration problems, low social skills, academic problems, and emotional dysregulation) interferes with a child's ability to adjust to the social demands of the classroom and makes it more likely these very problems will escalate over the course of the year without intervention. For instance, several studies examining variation in the impact of the Good Behavior Game, another universal prevention intervention targeting teacher use of effective classroom management practices, demonstrated that intervention effects were strongest among youth who displayed a high level of antisocial behavior at baseline (Kellam et al. 1998; van Lier et al. 2005), highlighting the importance of investigating the potential for baseline risk as a factor influencing variation of outcomes in prevention interventions. Therefore, consistent with a developmental epidemiological lens and prior research, we expected that students with more disruptive behavior or lower levels of social emotional competence and academic functioning would benefit more from the intervention than similar students in the control group (e.g., moderating effects of baseline risk on outcomes).

Methods

Participants and Setting

Participants in this group RCT were 105 teachers and 1817 students in kindergarten to third grade, from nine urban schools in a school district in the Midwestern part of the USA. All the schools were implementing school-wide Positive Behavior Supports and Interventions (PBIS; Sugai et al. 2002), with high fidelity (scores above 90% as gathered and reported by independent state evaluators). PBIS is a universal prevention intervention focused on teaching behavioral expectations and providing reinforcement to students for demonstrating those expectations. District-wide implementation of PBIS added strength to the design because it provided uniformity of behavior support programs across intervention and control schools. Nearly all schools across the country are implementing a multitude of social-emotional or behavioral intervention programs that often operate from different perspectives and levels of efficacy. For instance, on average, each school in a recent trial was implementing five separate social emotional intervention programs in a given year (Bradshaw et al. 2010). Therefore, having PBIS added consistency and reduced noise across buildings. However, even when PBIS is implemented with high fidelity at the school level, teachers may still struggle with proactive classroom management. For instance, in a study using data from classrooms in this trial, only three out of 34 teachers provided more positive than negative attention to students, with all but one teacher below the recommended 4:1 ratio of positive to negative attention (Reinke et al. 2013a, 2013b). Thus, offering a classroom-level intervention was expected to have an additive effect to school-wide PBIS.

As indicated in the participant flowchart (see Fig. 1), the study had high rates of enrollment for eligible teachers (96%) and students (84%). All teacher participants and parents of student participants provided written consent, and all students provided written assent to participate in the study. A blocked cluster random assignment design was utilized. Teachers were randomly assigned at the classroom level to receive IY TCM or to a wait list, business as usual control group with the constraint that the number of intervention teachers be no more than one more or less than the number of control teachers. Control teachers were offered the intervention 2 years following the year they participated. Teacher participants were recruited and randomized across three cohorts [year 1: 34 teachers (17 intervention), 577 students; year 2: 34 teachers (17 intervention), 571 students; year 3: 37 teachers (19 intervention), 670 students]. The majority of teacher participants were female (97%) and white (75%; 22% African American, and 3% other). The average years of teaching experience was 11 with a standard deviation of 8. The child sample included more males (52%) and African American students (76%; 22% White, and 2% other), 61% of the student sample qualified for free or reduced lunch, and 9% of the sample received special education services (see Table 1 for student demographics).

Procedures

The University Institutional Review Board and the participating school district approved the study protocol. Teachers and students were recruited at the beginning of the school year. Data were collected at the beginning of the school year, prior to the intervention, and at the end of the school year, post-intervention. All pre-intervention assessments occurred in early October. Post-intervention assessments were collected in late April and May of the same academic year.

Intervention Condition Teachers in the IY TCM condition attended three sets of two full-day group trainings in late October, December, and February. All trainings were co-facilitated by two doctoral-level IY TCM group leaders who were supervised by the program developer; one of these trainers also served as a coach.

IY TCM is a comprehensive curriculum for improving teacher classroom management skills. Specifically, teachers are trained to use proactive classroom management practices such as using behavior-specific praise, building positive relationships with students, using pre-corrective

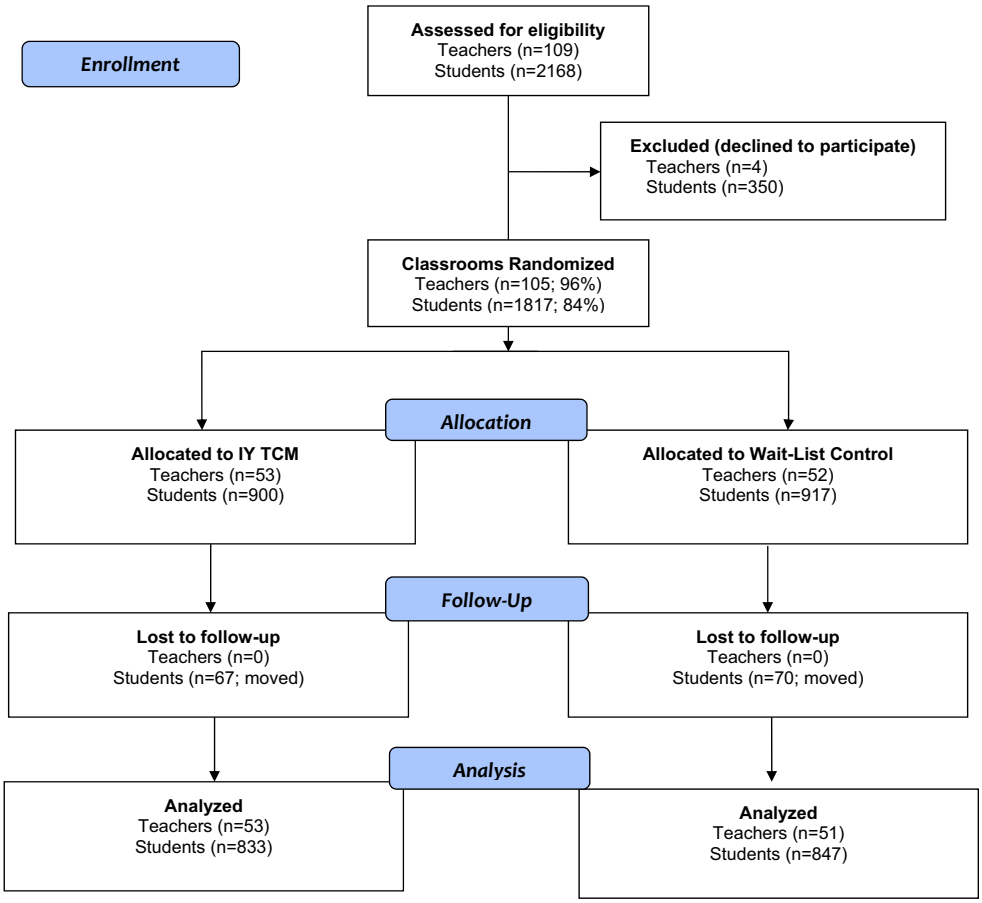

Fig. 1 IY TCM randomization participant flowchart

statements, teaching classroom expectations, and using proximity to reduce disruptive behavior. Much time during all training sessions is devoted to observing video vignettes of actual teacher interactions with students. The trainers prompt reflections with Socratic questioning about the videos ("What is the student learning?," "How would you respond in this situation?") and facilitate group discussion. These conversations spark role-plays to practice challenging interactions. In this way, teachers serve as models for others and/or get feedback about improving their skills.

Table 1 Student participant demographic information

	Control		Treatment	
	Mean	SD	Mean	SD
Age	7.06	1.09	7.15	1.22
% female	50		48	
% free or reduced lunch	61		60	
% special education	10		8	
% African American	74		76	
% White	23		22	
% other race	3		2	
% grade K	27		28	
% grade 1	28		27	
% grade 2	30		19	
% grade 3	15		26	

IY TCM Coaching The IY TCM coach was a doctoral-level special educator who was also a certified IY TCM group facilitator. The coaching model is learner-centered, supportive, and collaborative, and focuses on building on teachers' strengths (see Reinke et al. 2012). In between each workshop training session, the IY TCM coach observed the teachers in the classroom and met with them individually for up to 1 h on a weekly basis. The coach recorded any contact with teachers, including brief check-ins to reviewing strategies and schedule the next meeting. The overall mean time spent with a teacher by the coach outside of observing in the classroom was 28 min (range = 4 to 120 min). Time spent with the coach and number of coaching visits varied by teacher and was based on need for support. The overall mean number of coaching sessions for teachers in the sample was 7 (range = 3–12). Coaching was differentiated based on level of disruptive behavior and implementation of effective classroom management practices (see Reinke et al. 2012).

Fidelity of Implementation Fidelity of implementation of the IY TCM training and teacher implementation skills were monitored over the course of the year and are documented in a separate report (see Reinke, Herman, Stormont, & Newcomer 2014). In summary, teachers in the intervention were all exposed to the training; nearly all teachers attended all six training sessions (attendance rate was 94–100% for each training session) and the few teachers who missed a training due to illness or other reason met with the IY TCM coach to review missed material. The average amount of time each teacher spent with the coach over the course of the year was 6 h.

Teacher fidelity to use of effective proactive classroom management practices was monitored through direct observations which occurred in both IY TCM and control classrooms across four time points (October, December, February, and April). The results of a repeated measures ANOVA revealed a statistically significant difference between IY TCM and control teachers, indicating that IY TCM teachers used more proactive strategies over time in comparison to control teachers (Wilks' $\lambda = 0.89$, $F(3, 97) = 4.22$, $p < 0.01$, $\eta^2 = 0.12$) (see Fig. 2, available online along with a description of the measure). Further, because we were interested in whether there was an initial increase in proactive management implementation following teacher receipt of training and coaching, we conducted within subjects contrasts and found that time point one was significantly lower than time point 2 [$F(1, 50) = 44.99$, $p < 0.001$, $\eta^2 = 0.47$], time point 3 [$F(1, 50) = 52.57$, $p < 0.001$, $\eta^2 = 0.51$], and time point 4 [$F(1, 50) = 37.59$, $p < 0.001$, $\eta^2 = 0.43$], meaning that teacher implementation improved significantly after receiving the intervention and maintained over time.

Measures

Student Demographics Free and reduced lunch status (FRL), race, sex, and special education status were obtained from the school district for all participating students. Students were coded as 1 if they received FRL and 0 if not. Student sex was coded as 1 for male and 0 for female. Students receiving special education were coded as 1 and if not 0. For the purposes of this study, student race was coded as African American, White, or Other Race.

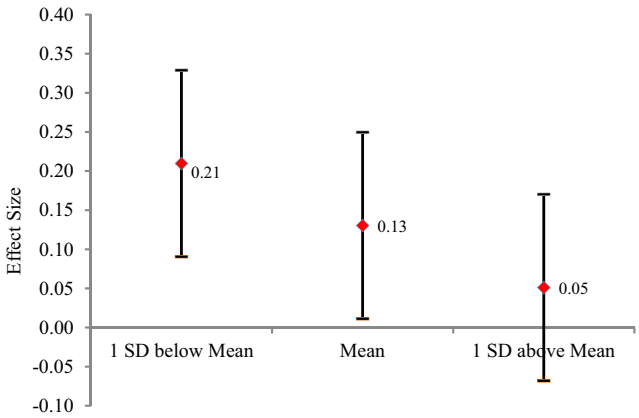

Fig. 2 Differential effect sizes and 95% confidence inervals on social competence varying by the pre-test

Teacher Report of Child Social Behavior and Academics The *Teacher Observation of Classroom Adaptation-Checklist* (TOCA-C; Koth et al. 2009) is a 54-item measure of child behavior. It was completed by the classroom teachers for each child. Teachers were asked to rate each child on the items referencing the past 3 weeks. The four subscales of the TOCA-C included in the present study were Disruptive Behaviors, Concentration Problems, Emotional Dysregulation, and Prosocial Behavior. The item responses ranged from 1 (never) to 6 (almost always). Prior studies support the factor structure of the TOCA-C (Koth et al. 2009) as well as strong evidence of subscale predictive validity. For instance, longitudinal data from the Prevention Intervention Research Center at Johns Hopkins University indicated that concentration problem scores in first grade predict likelihood of high school dropout and disruptive behavior scores in elementary school are strong predictors of violence in adolescence and adulthood (e.g., Petras et al. 2004). Prosocial behaviors, concentration problems, and disruptive behaviors also predict office discipline referrals (Pas et al. 2011). Previous research of the TOCA-C has found internal consistency estimates ranging from 0.86 to 0.96. For the current study, the internal consistency for each subscale ranged from 0.77 to 0.96.

The *Revised Social Competence Scale-Teacher version* (T-COMP; Gifford-Smith 2000) is a 17-item measure, which assesses the teacher's perception of a child's prosocial behavior, emotional self-regulation, and academic competence. Teachers were asked to rate each child in comparison to other children at their grade level. The total across all items provided an overall social competence score. For the purposes of this study, the academic competence subscale and the overall social competence scale were used. The item responses range from 0 (almost never) to 5 (almost always). The T-COMP scales have been shown to demonstrate strong internal consistency, have a consistent factor structure over time, and distinguish between high risk and normative samples (Gifford-Smith 2000). For the current study, the internal consistency for overall social competence ranged from 0.93 to 0.96 and from 0.92 to 0.93 for the academic competence subscale.

Standardized Academic Achievement In addition to teacher report of child academic competence, the *Woodcock-Johnson III* Normative Update Tests of Achievement (WJ III ACH; Woodcock, McGrew, & Mather 2007) was conducted

with each child. The WJ III is an assessment of child academic achievement. The present study included two subscales, Broad Reading and Broad Math. The WJ III ACH has strong psychometric properties (Woodcock, McGrew & Mather 2007; Bradley-Johnson, Morgan, & Nutkins 2004). Specifically, the test-retest reliability for the subtests included in the present study range from 0.80 to 0.95 (Bradley-Johnson et al. 2004). Research has also documented support for the construct validity of the Achievement subtests used in the current study (Bradley-Johnson et al. 2004). These data were gathered by trained independent researchers blind to the treatment status. Independent research assistants were trained in administration of the WJ III and required to pass a competency test before gathering data in the field.

Intervention Condition Dummy codes were used to indicate whether teachers received the intervention or wait list business as usual condition with intervention teachers being coded as 1 and control teachers coded as 0.

Analytic Approach

Missing Data

The original sample included 1817 students. Missing data occurred primarily on the outcome measures. The missing rates for the pre-tests of eight outcome measures ranged from 0.4 to 2.3% while the missing rates for the post-tests of eight outcome measures ranged from 6.4 to 7.3% in the overall sample. The maximum differential missing rates between the treatment and control groups are 2.7% for the pre-test and 0.70% for the post-test. The literature shows that when the outcome is included in the imputation model, there are very small differences between models that impute the outcome compared with those that do not (Kontopantelis et al. 2017). We chose to exclude the students whose post-tests were missing from the final analytic samples. The final analytic samples included nine schools (105 teachers and 1680 students for the analyses of social and behavioral outcomes; 105 teachers and 1685 students for the analyses of academic achievement outcomes). The maximum data missing rate in the final analytic samples was 1.8%. Multiple imputation using a Markov chain Monte Carlo (MCMC) method in SAS PROC MI was used to impute missing data on pre-test and other covariates and the outcome measures were included in the imputation models. We imputed five times given the small missing rate (Rubin 1987; Schafer and Olsen 1998).

Analysis of Main Effects

For each of the five imputed datasets, three-level hierarchical linear models (HLM), in which students (level 1) are nested within teachers (level 2) and teachers are nested within schools (level 3), were conducted using SAS PROC MIXED to examine the overall treatment effects student behavior and academic outcomes. Each student's pre-test and demographic information were included at level 1, and the treatment variable was at level 2 and its coefficient was assumed constant across level 3. SAS PROC MIANALYZE was used to combine the results from the analyses of five datasets. The statistical model is below:

$$\text{Level 1(student)}: \quad Y_{ijk} = \alpha_{0jk} + \sum_{q=1}^{Q} \alpha_{qjk} X_{qijk} + e_{ijk}, \quad e_{ijk} \sim N\left(0, \sigma^2\right)$$

Level 2(class)

$$: \quad \begin{array}{l} \alpha_{0jk} = \beta_{00k} + \beta_{01k}(Condition)_{jk} + \mu_{jk} \\ \alpha_{qjk} = \beta_{q0k}, q = 1, \ldots, Q. \end{array}, u_{jk} \sim N\left(0, \tau_2^2\right)$$

Level 3(school)

$$: \quad \begin{array}{l} \beta_{00k} = \gamma_{000} + \xi_k \\ \beta_{01k} = \gamma_{001} \\ \beta_{q0k} = \gamma_{q00}, q = 1, \ldots, Q. \end{array}, \xi_k \sim N\left(0, \tau_3^2\right)$$

where X_{qijk} represents student-level covariates, which include pre-test, age at pre-test, gender, race, FRL, special education status, grade level, and cohort year in the study. $(Condition)_{jk}$ is a binary variable indicating treatment condition (Condition = 0 for control group and Condition = 1 for treatment group). The parameter, γ_{001}, estimates the overall treatment effect, which is assumed constant across schools. σ^2, τ_2^2, and τ_3^2 are variance components for level 1, level 2, and level 3 conditional on these covariates.

Moderation Analysis

Moderation analyses were conducted to examine whether the treatment effects on child outcomes differed by demographic variables and pre-test measures of social emotional and disruptive behavior and academics. The statistical models were similar to the models with main effects but had the additional interaction term of treatment and moderator. Specifically, in the level 2 model, we included the treatment condition to predict to the coefficient of the pre-test that was grand mean centered.

$$\text{Level 2}(class): \quad \begin{array}{l} \alpha_{0jk} = \beta_{00k} + \beta_{01k}(Condition)_{jk} + \mu_{jk} \\ \alpha_{qjk} = \beta_{q0k} + \beta_{q1k}(Condition)_{jk}, q = 1, \ldots, Q. \end{array}, u_{jk} \sim N\left(0, \tau_2^2\right)$$

The parameter, β_{q1k}, estimates the moderator effects of the qth covariate and is assumed constant across schools ($\beta_{q1k} = \gamma_{q10}$).

Results

Descriptive Statistics

ESM 1 provides descriptive statistics for the study variables (available online). The effect sizes indicated that the baseline measures were equivalent between the two conditions. HLM results for three-level models examining the effects of IY TCM on social behavior and academic outcomes and moderation analyses are provided below. The unconditional intraclass correlations (ICCs) ranged from 0.02 to 0.05 at school level and from 0.06 to 0.18 at class level for social behavior measures. This suggests that bigger proportion of variance on the social behavior measures exist between-classrooms than between-schools and children with behavior problems may be clustered into the same classroom. The unconditional ICCs ranged from 0.03 to 0.07 at school level and from 0.04 to 0.08 at class level for academic outcome measure.

Main Effects on Social Behavior

Table 2 provides the main effects of the intervention on social behavior outcomes. In Table 2, the list of baseline variables included in the analyses is noted on the left-hand side. The outcome variables are indicated across the top of the table. Teacher ratings of African American students and boys indicated that they had more concentration problems, disruptive behavior, emotion dysregulation, and lower prosocial behavior and social competence. Further, students in special education were rated as having more concentration problems, and students receiving FRL had more concentration problem and lower social competence. These demographic variables were included as covariates in each model given the strong associations found in the literature between these variables and social behavior. Baseline measures of the same outcome variables were also covariates in the model. Therefore, the post-intervention scores represent change between baseline and post-assessment. Main effect analyses demonstrated that students in IY TCM showed significant improvement on teacher-reported prosocial behavior ($b = 0.14$, $p = 0.038$, $d = 0.13$), overall social competence ($b = 0.14$, $p = 0.032$, $d = 0.13$), and reduced emotional dysregulation ($b = -0.15$, $p < 0.001$, $d = -0.14$) compared to students in the control condition. There were no significant effects for disruptive behavior or concentration problems.

Main Effects on Academic Outcomes

Table 3 provides the main effects of the intervention on academic outcomes. Several demographic variables that are known to be associated with academic outcomes were included in the models as covariates. Students receiving special education services and African American students had lower teacher ratings of academic competence and lower reading and math scores. Younger students also had lower scores in reading and math. Teachers rated boys as having higher academic competence than girls. There were no main effects on standardized reading or math scores. However, teacher-reported academic competence demonstrated a trend favoring students in the IY TCM classroom in comparison to control classrooms ($b = 0.11$, $p = 0.074$, $d = 0.08$).

Moderating Effects

Tables 2 and 3 also provide the moderation effects of pre-test measures on outcomes. Demographic information as moderators were not statistically significant at an alpha of 0.05. Thus, only the results for pre-tests as moderators were reported here. Pretests were centered at the grand mean. Two of the pre-test measures had statistically significant moderation effects. Overall social competence was moderated by pre-test levels of social competence ($b = -0.09$, $p = 0.04$), indicating that students with lower social competence at pre-test who were in intervention classrooms demonstrated greater increases in social competence over time in comparison to similar peers in control classrooms. Similarly, students with lower levels of academic competence at pre-test demonstrated greater increases in academic competence over time in comparison to similar peers in the control classrooms ($b = -0.08$, $p < 0.001$). Figures 2 and 3 present the effect sizes and their 95% confidence intervals at the mean of the baseline pretest, and one standard deviation below and above the mean pre-test for the social competence and academic competence measures, respectively.

Discussion

This group RCT investigated the efficacy of the IY TCM as a stand-alone program among teachers in classrooms kindergarten to third grade on child social behavior and academic outcomes. It was hypothesized that students in the classrooms of teachers who received the IY TCM would exhibit reductions in disruptive behavior, concentration problems, and emotional dysregulation. In addition, students in the IY TCM classrooms were expected to demonstrate increases in prosocial behavior, social competence, and improvements in academics.

The first hypothesis was partially supported in that students in IY TCM classrooms demonstrated a significant reduction in emotional dysregulation and improvements in prosocial behavior and social competence relative to students in business as usual classrooms. However, there were no significant findings for disruptive behavior or concentration problems. It was expected that IY TCM would demonstrate significant proximal impacts on concentration difficulties and disruptive

Table 2 HLM results for three-level model examining the effects of IY TCM on social emotional and disruptive behaviors

	Concentration problems			Disruptive behavior			Prosocial behavior			Emotional dysregulation			Social competence		
Variable	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*
Intercept	2.89**	0.47	< 0.01	1.81**	0.25	< 0.01	4.63**	0.28	< 0.01	2.39**	0.38	< 0.01	3.51**	0.23	< 0.01
Age	− 0.05	0.09	0.54	0.01	0.04	0.85	− 0.01	0.05	0.90	− 0.02	0.07	0.75	− 0.02	0.04	0.62
Sex	− 0.52**	0.12	< 0.01	− 0.11**	0.02	< 0.01	0.12**	0.04	< 0.01	− 0.20**	0.03	< 0.01	0.17**	0.03	< 0.01
Lunch status	0.22*	0.10	0.02	0.02	0.02	0.49	− 0.05	0.05	0.24	0.07	0.04	0.11	− 0.07*	0.03	0.03
Special education	0.40**	0.11	< 0.01	− 0.03	0.05	0.56	− 0.07	0.06	0.23	0.11	0.07	0.158	− 0.06	0.06	0.32
African American	0.40**	0.10	< 0.01	0.12**	0.03	< 0.01	− 0.16**	0.03	< 0.01	0.19**	0.05	< 0.01	− 0.18**	0.04	< 0.01
Other race	0.15	0.12	0.20	0.03	0.07	0.64	0.04	0.05	0.36	− 0.17	0.10	0.10	0.11*	0.05	0.02
Year 2	− 0.31	0.32	0.33	0.04	0.05	0.42	0.52**	0.07	< 0.01	− 0.07	0.06	0.27	0.30**	0.09	< 0.01
Year 3	0.24	0.13	0.07	0.08*	0.03	0.03	− 0.04	0.07	0.56	0.15*	0.08	0.07	− 0.08	0.06	0.18
Grade 1	− 0.09	0.16	0.60	− 0.04	0.06	0.47	0.09	0.11	0.42	− 0.10	0.08	0.19	0.08	0.08	0.35
Grade 2	− 0.17	0.16	0.26	− 0.07	0.10	0.46	0.18	0.15	0.21	− 0.12	0.14	0.39	0.20	0.12	0.09
Grade 3	0.00	0.22	0.99	− 0.08	0.12	0.50	− 0.01	0.18	0.95	0.00	0.20	0.99	0.01	0.12	0.93
Pre-test	0.37	0.21	0.08	0.81**	0.04	< 0.01	0.82**	0.04	< 0.01	0.80**	0.03	< 0.01	0.89**	0.04	< 0.01
Intervention	− 0.07	0.08	0.38	− 0.04	0.05	0.43	0.14*	0.07	0.04	− 0.15**	0.04	< 0.01	0.14*	0.06	0.03
Step 2: moderation analysis															
Pre-test × intervention	0.03	0.05	0.59	− 0.08	0.06	0.21	− 0.04	0.04	0.38	− 0.08	0.05	0.09	− 0.09*	0.04	0.04

Year = cohort; year 1 is the reference group for cohort; kindergarten is the reference group for grade

*$p < 0.05$; **$p < 0.01$ or less

Table 3 HLM results for three-level model examining the effects of IY TCM on academic outcomes

	Academic competence			Reading			Math		
Variable	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*
Intercept	3.72**	0.25	< 0.01	119.15**	4.06	< 0.01	118.03**	2.64	< 0.01
Age	− 0.01	0.04	0.76	− 2.42**	0.68	<0.01	− 3.58**	0.29	< 0.01
Sex	0.16**	0.03	< 0.01	0.57	0.33	0.09	− 0.31	0.55	0.58
Lunch status	− 0.05	0.04	0.27	− 0.66	0.35	0.06	− 0.60	0.41	0.15
Special education	− 0.14*	0.06	0.02	− 2.65**	0.59	< 0.01	− 3.96**	1.27	< 0.01
African American	− 0.25**	0.03	< 0.01	− 1.74**	0.28	< 0.01	− 2.18**	0.43	< 0.01
Other race	− 0.03	0.10	0.75	− 0.80	0.52	0.13	1.42	1.48	0.34
Year 2	− 0.12	0.06	0.07	− 0.10	0.65	0.88	3.01	1.50	0.04
Year 3	− 0.04	0.04	0.24	− 0.64	0.57	0.26	0.29	1.72	0.87
Grade 1	− 0.02	0.10	0.87	0.50	1.37	0.72	5.38**	1.30	< 0.01
Grade 2	0.12	0.13	0.36	− 0.59	1.46	0.69	7.63**	1.19	< 0.01
Grade 3	− 0.09	0.12	0.47	1.19	2.02	0.56	12.38**	0.84	< 0.01
Pre-test	0.83**	0.02	< 0.01	0.79**	0.03	< 0.01	0.75**	0.02	< 0.01
Intervention	0.11	0.06	0.08	− 0.15	0.23	0.53	− 0.03	0.39	0.94
Step 2: moderation analysis									
Pre-test × intervention	− 0.08**	0.02	< 0.01	− 0.04	0.03	0.14	0.01	0.04	0.72

Year = cohort; year 1 is the reference group for cohort; kindergarten is the reference group for grade

*$p < 0.05$; **$p < 0.01$ or less

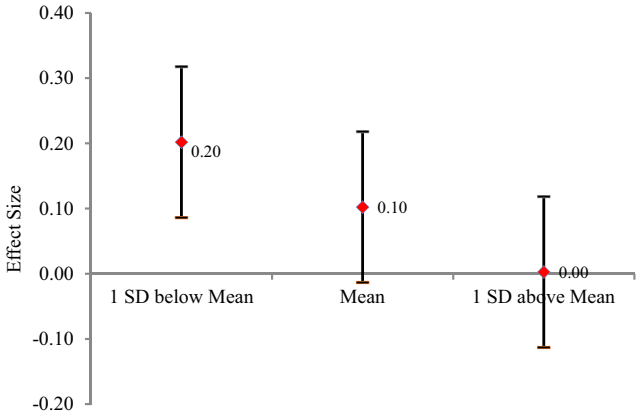

Fig. 3 Differential effect sizes and 95% confidence inervals on academic competence varying by the pre-test

behavior such as has been found with other universal behavior interventions (Bradshaw et al. 2012; Kellam et al. 2008). Notably, all schools in the present study were implementing a school-wide behavior support program (PBIS) with high fidelity. Prior studies have found significant effects of these school-wide practices in reducing disruptive behaviors. Thus, in this study, IY TCM may have had smaller effects on disruptive behaviors above and beyond any changes that may have been related to implementing PBIS.

IY TCM helped to decrease emotion dysregulation. Students with emotion regulation skills are more likely to be academically successful (Raver et al. 2017). Thus, decreasing dysregulation can be a proximal outcome toward improving student achievement. In addition, students in the intervention demonstrated improvements in prosocial behaviors and social competence. The finding that the intervention improved social competence at the end of the year for youth with lower levels of baseline social competence is particularly encouraging. This demonstrates that a universal classroom management approach can help to mitigate risk for youth with initially lower levels of social competence potentially leading to reductions in the number of students who would otherwise need more intensive supports in this area.

The absence of intervention effects on academic achievement scores were not entirely surprising given that performance on standardized measures of achievement is a fairly distal outcome relative to changes in classroom behaviors. Improving academic enablers, such as coaching academic persistence, promoting effective problem solving and social skills (Elliott et al. 2004), and increasing emotional regulation skills (Zimmerman 1998), may increase student access to instruction and learning. Here we found evidence that IY TCM promoted teacher-rated academic competence for youth with lower levels of competence at the start of the year.

The effect sizes on self-regulation, prosocial skills, and social competence were relatively modest; however, small effect sizes are common in longitudinal universal prevention studies (Durlak et al. 2011). Small effects are expected from universal preventive interventions given that they are delivered to entire populations with varying degrees of risk (i.e., many individuals would not develop behavior or academic problems even without the intervention); yet very small effects on a population level can result in dramatic improvements in public health outcomes (NRC and IOM 2009). Prior studies of the IY TCM as a stand-alone intervention systematically either included only students at risk for conduct problems (Hitchings et al. 2013) or had a larger percentage of students with elevated conduct problem scores (Hutchings et al. 2007). Thus, there was more room for improvement among these samples than in the current study, explaining the larger effect sizes seen in those studies.

Study Limitations

While the findings are interesting and important, this study is not without some limitations. First, all schools in this trial were implementing PBIS with high fidelity. It is possible that this school-wide universal intervention to support student positive behavior reduced the effect that could be detected with the addition of a classroom level intervention, particularly on student disruptive behavior. Second, the findings from the study were predominantly on measures of teacher report. Teachers were also the recipients of the training to implement IY TCM practices, leading to the potential that teachers who received training may have rated their students as improved due to being exposed to the intervention. Despite this, teachers are the most common source of information used to assess social behavior and determine special education evaluations (Zima, Haltburt, Kinapp, et al. 2005); thus, their ratings are important in the context of school-based interventions and have been shown to predict social behavioral problems (Koth et al. 2009; Reinke et al. 2008; Schaeffer et al. 2003). Importantly, independent observations revealed that teachers who received the IY TCM demonstrate significant improvement in their use of proactive classroom management strategies in comparison to control teachers. The lack of long-term follow-up in this study is also a limitation. Some of the more proximal findings in this study may lead to improvements in more distal outcomes, but this cannot be confirmed without follow-up. Lastly, in this study, we do not report the indirect impacts of teacher behavior on student outcomes. Given the complex nature of the IY TCM intervention and the need to systematically investigate the potential mechanisms for mediation, we plan to conduct these analyses in a separate study in hopes to elucidate the proximal mechanisms of change leading to student outcomes.

Future Directions

The present findings suggest that IY TCM holds promise as a universal prevention program for supporting teacher use of effective classroom management practices and for promoting youth prosocial behavior, emotion regulation, and social competence. Future research will examine the mechanisms by which IY TCM impacts youth outcomes. Additionally, given the intense and comprehensive nature of the IY TCM, researchers may want to look toward identifying the critical ingredients of the intervention by dismantling the IY TCM components. Determining whether particular components of the training have a greater impact on particular social or academic behaviors could increase feasibility and focus for schools. Further, investigating whether some teachers benefit more from the IY TCM training than others would be of interest. Perhaps fewer training days are necessary for teachers who have a higher rate of proactive classroom management practices. Additionally, future research could systematically assess additional academic enablers that could lead to later academic achievement improvements for students. For instance, the finding that IY TCM impacted academic competence for those students with lower levels of academic competence at baseline raises the question whether these students will demonstrate further improvements in academic achievement more distally. Long-term follow-up of students may help to specify the connection between these more proximal indicators on later improvement in academic achievement. Finally, some students have co-occurring problems, including academic and behavioral issues. Future research can use person-centered methods, such as latent profile analysis, to determine if students with co-occurring problems differentially respond to this intervention.

Funding This study was supported by a 5-year grant from the US Department of Education, Institute of Education Sciences (no. R305A100342; to the first and second authors) submitted to CFDA 84.324A.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

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

References

Bandura, A. (1977). *Social learning theory*. Oxford England: Prentice-Hall.

Bradley-Johnson, S., Morgan, S. K., & Nutkins, C. (2004). Review of The Woodcock-Johnson Tests of Achievement: Third Edition. *Journal of Psychoeducational Assessment, 22*(3), 261-274. https://doi.org/10.1177/073428290402200305.

Bradshaw, C., Mitchell, M., & Leaf, P. (2010). Examining the effects of schoolwide positive behavior interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*, 133–148. https://doi.org/10.1177/1098300709334798.

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics, 130*(5), e1136–e1145. https://doi.org/10.1542/peds.2012-0243.

Bradshaw, C. P., Waasdorp, T., & Leaf, P. (2015). Examining the variation in the impact of school-wide positive behavior interventions and supports: Finding from a randomized controlled effectiveness trial. *Journal of Educational Psychology, 2*, 546–557.

Conroy, M., Sutherland, R., Haydon, Stormont, M., & Hardon. (2009). Preventing and ameliorating young children's chronic problem behaviors: An ecological classroom-based approach. *Psychology in the Schools, 46*, 3–17.

Curby, T. W., Rimm-Kaufman, S. E., & Abry, T. (2013). Do emotional support and classroom organization earlier in the year set the stage for higher quality instruction? *Journal of School Psychology, 51*, 557–569.

Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.

Elliott, S. N., DiPerna, J. C., Mroch, A. A., & Lang, S. C. (2004). Prevalence and patterns of academic enabling behaviors: An analysis of teachers' and students' ratings for a national sample of learners. *School Psychology Review, 33*, 297–304.

Farrell, A. D., Henry, D. B., & Bettencourt, A. (2013). Methodological challenges examining subgroup differences: Examples from universal school-based youth violence prevention trials. *Prevention Science, 14*, 121–133. https://doi.org/10.1007/s11121-011-0200-2.

Gifford-Smith, M. (2000). *Teacher social competence scale, fast track project technical report*. Durham, NC: Duke University.

Hitchings, J. Martin-Forbes, P., Daley, D., & Williams, M. (2013). A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. *Journal of School Psychology, 51*, 571–585.

Hutchings, J., Daley, D., Jones, K., Martin, P., Bywater, T., and Gwyn, R. (2007). Early results from developing and researching the Webster-Stratton Incredible Years Teacher Classroom Management Training Programme in North West Wales. *School of Psychology*, University of Wales, 1–29.

Ialongo, N., Poduska, J., Werthamer, & Kellam. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders, 9*, 146–160.

Jones, V. F., & Jones, L. S. (2004). *Comprehensive classroom management, creating communities of support and solving problems* (7th ed.). Boston: Allyn & Bacon.

Kellam, S. G., Werthamer-Larsson, L., Dolan, L. J., Brown, C. H., Mayer, L. S., Rebok, G. W., et al. (1991). Developmental epidemiologically based preventive trials: Baseline modeling of early target behaviors and depressive symptoms. *American Journal of Community Psychology, 19*(4), 563–584.

Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology, 10*, 165–185.

Kellam, S. G., Brown, C. H., Poduska, J., Ialongo, N., Wang, W., Toyinbo, P., et al. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95*(Suppl 1), S5–S28.

Kontopantelis, E., White, I. R., Sperrin, M. & Buchan, L. (2017). Outcome-sensitive multiple imputation: A simulation study. *BMC Medical Research Methodology, 17* (2). DOI 10.1186/s12874-016-0281-5.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2009). Teacher observation of classroom adaptation-checklist: Development and factor structure. *Measurement and Evaluation in Counseling and Development, 42*, 15–30.

National Research Council. (2002). *Minority students in special and gifted education. Committee on Minority Representation in Special Education*. Washington DC: National Academy Press.

National Research Council and Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.

Pas, E. T., Bradshaw, C. P., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools, 48*, 541–555.

Petras, H., Chilcoat, H., Leaf, P., Ialongo, N., & Kellam, S. (2004). The utility of teacher ratings of aggression during the elementary school years in identifying later violence in adolescent males. *Journal of the American Academy of Child and Adolescent Psychiatry, 1*, 88–96.

Raver, C., Adams, K., & Blair, C. (2017). Self-regulation in early childhood: Implications for motivation and achievement. In A. Elliott, C. Dweck, & D. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (pp. 408–430). New York, NY: Guilford Press.

Reinke, W. M. & Herman, K. C. (2002). Creating school environments that deter antisocial behaviors in youth. *Psychology in the Schools, 39*, 549–559.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A classwide consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review, 37*, 315–332.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly, 26*, 1–13.

Reinke, W. M., Stormont, M., Webster-Stratton, C., Newcomer, L., & Herman, K. C. (2012). The incredible years teacher training: Using coaching to support generalization to real world classroom settings. *Psychology in the Schools, 49*, 416–428.

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

Reinke, W. M., Herman, K. C., Stormont, M., Newcomer, L., & David, K. (2013b). Illustrating the multiple facets and levels of fidelity of implementation to a teacher classroom management intervention. *Administration and Policy in Mental Health and Mental Health Services Research, 40*, 494–506.

Reinke, W. M. Stormont, M., Herman, K. C., & Newcomer, L. (2014). Using coaching to support teacher implementation of classroom-based interventions. *Journal of Behavioral Education, 23*, 150–167.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc..

Schaeffer, C. M., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology, 39*, 1020–1035.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33*(4), 545–571. https://doi.org/10.1207/s15327906mbr3304_5.

Schochet, P. Z., Puma, M., & Deke, J. (2014). *Understanding variation in treatment effects in education impact evaluations: An overview of quantitative methods* (NCEE Report 2014–4017). Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children, 31*, 351–380. https://doi.org/10.1353/etc.0.0007.

Sugai, G., Horner, R., & Gresham, F. (2002). Behaviorally effective school environments. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Prevention and remedial approaches* (pp. 315–350). Bethesda, MD: NASP.

van Lier, P. A. C., Vuijk, P., & Crijnen, A. A. M. (2005). Understanding mechanisms of change in the development of antisocial behavior: The impact of a universal intervention. *Journal of Abnormal Child Psychology, 33*, 521–535. https://doi.org/10.1007/s10802-005-6735-7.

Walker, H. M., Colvin, G., & Ramsey, X. (1995). *Antisocial behavior in school: Strategies and best practices*. Belmont, CA: Thomson Brooks/Cole Publishing Co..

Webster-Stratton, C. & Reid, J. (2004). Strengthening social and emotional competence in young children-the foundation for early school readiness and success: Incredible Years Classroom Social Skills and Problem-Solving Curriculum. *Infants & Young Children, 17*, 96–113.

Webster-Stratton, C., Reid, M., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in head start. *Journal of Clinical Child Psychology, 30*, 283–302.

Webster-Stratton, C., Reid, M., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33*, 105–124.

Webster-Stratton, C., Reid, J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49*, 471–488.

Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. (2011). The Incredible Years Teacher Classroom Management Training: The methods and principles that support fidelity of training delivery. *School Psychology Review, 40*, 509–529.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside.

Zima, B. T., Hurlburt, M. S., Knapp, P., et al. (2005). Quality of publicly-funded outpatient specialty mental health care for common childhood psychiatric disorders in California. *Journal of American Academy of Child and Adolescent Psychiatry, 44*(2), 130–144.

Zimmerman, B. J. (1998). *Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models*. In D. H. Schunk & B. J. Zimmerman (Eds.), Self-regulated learning: From teaching to self-reflective practice (pp. 1–19). New York, NY, US: Guilford Publications.