



Investigating teacher and student effects of the Incredible Years Classroom Management Program in early elementary school[☆]



Desiree W. Murray^{a,*,1}, David L. Rabiner^b, Laura Kuhn^c, Yi Pan^{c,2}, Raha Forooz Sabet^d

^a Duke University and Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill, United States

^b Duke University, United States

^c Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill, United States

^d University of Miami, United States

ARTICLE INFO

Action Editor: Steve Kilgus

Keywords:

Incredible Years
Classroom climate
Teacher training
Rural schools
Social-emotional

ABSTRACT

The present paper reports on the results of a cluster randomized trial of the Incredible Years[®] Teacher Classroom Management Program (IY-TCM) and its effects on early elementary teachers' management strategies, classroom climate, and students' emotion regulation, attention, and academic competence. IY-TCM was implemented in 11 rural and semi-rural schools with K-2 teachers and a diverse student sample. Outcomes were compared for 45 teachers who participated in five full day training workshops and brief classroom consultation and 46 control teachers; these 91 teachers had a total of 1192 students. A high level of teacher satisfaction was found and specific aspects of the training considered most valuable for early elementary teachers were identified. Hierarchical linear modeling indicated a statistically significant intervention effect on Positive Climate in the classroom ($d = 0.45$) that did not sustain into the next school year. No main effects on student outcomes were observed, although a priori moderator analyses indicated that students with elevated social-behavioral difficulties benefitted with regard to prosocial behavior ($d = 0.54$) and inattention ($d = -0.34$). Results highlight potential benefits and limitations of a universal teacher training program for elementary students, and suggest strategies for future delivery of the IY-TCM program and areas for future research.

The importance of teachers' classroom management skills is well-established for supporting young children's social and behavioral competence (Hamre & Pianta, 2005; Matheson & Shriver, 2005; Pianta, La Paro, Payne, Cox, & Bradley, 2002). Classroom management typically refers to the manner in which teachers implement and maintain order and discipline with students and the structure and supervision that affect student opportunities to learn (Oliver, Wehby, & Reschly, 2011). Well-managed classrooms are characterized by teachers' ability to monitor student attention and performance, establish behavioral expectations, and consistently

[☆] Portions of the data examined were presented as a poster at the 2014 meeting of the Society for Research on Educational Effectiveness (SREE). Washington, D.C. The research reported here was supported by the U.S. Department of Education, Institute of Education Sciences through Grant R305A090361, awarded and administered through Duke University and also supported by Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill. The contents of this publication are those of the authors and do not represent views or policies of the Institute of Education Sciences. We graciously acknowledge the many teachers, schools, and research staff who made this study possible.

* Corresponding author at: Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill, Campus Box 8180, 105 Smith Level Road, Chapel Hill, NC 27599-8180, United States.

E-mail address: desiree.murray@unc.edu (D.W. Murray).

¹ Dr. Desiree W. Murray is a trained mentor in the Incredible Years Teacher Classroom Management Program, and receives compensation from community organizations for providing trainings and consultation.

² Deceased April 15, 2016.

implement rules and procedures that prevent problems from occurring (Oliver et al., 2011). Well-trained and supportive teachers who use high rates of praise, proactive teaching strategies, and non-harsh discipline foster the development of social and emotional skills that can also facilitate learning (Webster-Stratton, Reid, & Stoolmiller, 2008). Indeed, children in well-managed classrooms also spend more time engaged in academic tasks, progress at a more rapid pace, and demonstrate higher levels of academic achievement (Cameron, Connor, & Morrison, 2005; Connor, Morrison, & Slominski, 2006; Konstantopoulos & Chung, 2011; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008).

In addition to benefits from specific teacher management skills, classroom climate also appears to impact young students' social-emotional, cognitive, and academic development (Hamre & Pianta, 2010; Jennings & Greenberg, 2009). Climate includes teacher-student relationships, which have become recognized as a critical factor of the classroom environment (Downer, Sabol, & Hamre, 2010). In particular, the quality of emotional support (i.e., closeness, warmth, and positivity) that students receive from teachers has been shown to be related to their motivation and engagement, which predicts peer acceptance (Silver, Measelle, Armstrong, & Essex, 2005), as well as academic and behavioral outcomes several years later (Hamre & Pianta, 2001). Thus, there is considerable support for enhancing teachers' classroom management skills and strengthening their relationships with students. The present study reports on the evaluation of one program that addresses both of these goals - the Incredible Years® Teacher Classroom Management Program (IY-TCM).

1. Incredible Years® Teacher Classroom Management Program

The Incredible Years® Teacher Classroom Management Program (IY-TCM) is part of a comprehensive series of interventions including parent, child, and teacher training components that were designed to prevent and treat aggressive behavior and conduct problems in young children (Webster-Stratton, 1994). IY-TCM is provided in 6 monthly full-day workshops with two trained co-leaders and 12–15 teachers in each group. Its overall content addresses teacher management skills and classroom climate, with specific topics covering the following: building positive relationships with students and parents; proactive classroom management strategies (e.g., instructions, rules, transitions); effective use of incentives; “coaching” students' social, academic and emotional development and ability to persist on frustrating and challenging tasks, teaching calm-down and problem-solving; differential attention; and positive discipline techniques such as redirection, ignoring, and time out (Webster-Stratton, 2012). IY-TCM incorporates active learning training methods recommended for teachers (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009), including video-modeling, behavioral rehearsal of key skills through role plays, classroom practice assignments, and teacher goal setting and self-monitoring (Webster-Stratton, Reinke, Herman, & Newcomer, 2011). Peer support strategies are embedded in professional learning communities to support sustainable practice changes.

Among universal teacher classroom management interventions for young children, IY-TCM has substantial evidence of support, including eight randomized controlled trials. Furthermore, studies have been conducted by both the developer (Webster-Stratton et al., 2008; Webster-Stratton, Reid, & Hammond, 2001; Webster-Stratton, Reid, & Hammond, 2004), as well as several independent investigators (Baker-Henningham, Walker, Powell, & Gardner, 2009; Hutchings, Martin-Forbes, Daley, & Williams, 2013; McGilloway et al., 2010; Morris et al., 2014; Raver et al., 2009). Across studies, results consistently demonstrate benefits on both observational and self-report measures of harsh/critical and warm/responsive behavior, use of praise and social-emotional instruction, and effectiveness of commands and discipline strategies. Improvements are typically seen in students' emotion regulation, social skills and social problem-solving, cooperation, and problem behaviors including aggression, with larger effects observed for children at higher risk levels.

Despite these encouraging findings, there are several limitations to the existing IY-TCM research that will be addressed in the present study. First, previous studies have generally included intervention components in addition to IY-TCM, such as a parenting program, child curriculum, or intensive classroom coaching (Morris et al., 2014; Raver et al., 2009). In fact, only one published randomized controlled trial (RCT) examined IY-TCM as an independent intervention without coaching (Hutchings et al., 2013), with positive effects seen on teacher and child behaviors in a sample of children aged 3–7 years (average age of 4.8 years). More specifically, the Hutchings study showed significant effects on teacher positive behaviors and commands (but not negative teacher behaviors) and on three of six observed child classroom behaviors. Second, the majority of participants in prior studies were preschool teachers, who may differ from early elementary teachers in important ways. For instance, they might possess fewer classroom management skills (Pianta, La Paro, & Hamre, 2008) and have less experience with positive behavior support systems (Benedict, Horner, & Squires, 2007). Relatedly, the majority of students examined in prior work appear to come from adverse backgrounds (e.g., high poverty schools in urban areas) or have clearly identified risk behaviors. Whether IY-TCM may provide comparable benefits to lower risk, older populations is unknown.

Given the novel implementation of IY-TCM in this study, it is important to consider its acceptability for early elementary teachers working in diverse schools with students who are both at-risk and low risk. There are several aspects of IY-TCM which may enhance acceptability, including its collaborative approach, use of a group format which facilitates teacher peer support and learning, and its flexibility in addressing teacher concerns and needs (Webster-Stratton et al., 2011). Prior research on IY-TCM has indicated that teachers perceive the strategies taught to be useful (Carlson, Tired, Bender, & Benson, 2011; Williford & Shelton, 2008), although this is based on preschool teachers working in low-income settings. It is certainly possible that early elementary teachers may have different experiences or perceptions related to differences in pre-service training or in increased accountability for academic outcomes at the elementary level. Indeed, prior evaluation of IY-TCM in combination with the IY classroom skills curriculum indicated that Head Start teachers needed a higher level of training support than did kindergarten and first grade teachers (Webster-Stratton et al., 2008). Should IY-TCM prove efficacious for the population targeted in this study, information on its acceptability would be important for future dissemination efforts.

1.1. Self-regulation as a framework for evaluating student outcomes

Teachers' use of more effective management strategies and social-emotional supports may not only externally regulate student behavior in the classroom, but may also help students learn to regulate their own behavior. Self-regulation is considered a particularly useful framework for understanding the underlying mechanisms by which children gain skills that translate into positive behavioral and social-emotional outcomes (Murray, Rosanbalm, & Christopoulos, 2016). From an applied perspective, self-regulation can be defined as the act of managing cognition and emotion to enable goal-directed actions, and can be categorized into cognitive and emotional regulation (Murray, Rosanbalm, Christopoulos, & Hamoudi, 2015). An important indicator of cognitive regulation is attention, which is foundational for higher-order thinking and appears to uniquely predict academic achievement (Rabiner, Coie, & Conduct Problems Prevention Research Group, 2000). Self-regulation strategies also enable children to modify their behavior in ways that match classroom expectations and rules, inhibit inappropriate, impulsive, and aggressive behaviors (i.e., response inhibition), and organize behavior to complete tasks (Murray et al., 2016). Managing frustration and other strong, negative emotions, and solving social problems are also important for social-emotional adjustment within schools (Eisenberg & Sulik, 2012).

Theoretically, effective teacher classroom management and warm, responsive student-teacher relationships promoted by the IY-TCM program should enhance skills in each self-regulation domain investigated in the present study: attention, emotion regulation, and prosocial behavior. By enhancing classroom structure, consistency, and teachers' use of positive behavior management strategies, students may improve their inhibitory control in managing frustration and delaying gratification (e.g., emotion regulation) and may be better able to attend to their academic work. IY-TCM's focus on building positive relationships with students may also help students' abilities to understand, express, and modulate their thoughts, feelings, and behavior in a process described as "co-regulation" (Murray et al., 2015). In addition, IY-TCM directly addresses academic competence through academic and persistence "coaching," in which teachers provide descriptive commenting on individually-targeted academic skills and behaviors such as accurately sounding out words, performing math calculations, considering alternative solutions to problems, working carefully, and paying attention. In particular, coaching persistence in the face of challenging or frustrating tasks may translate into increased academic engagement (Webster-Stratton, 2012).

Evaluation of whether teacher behavior management practices and social-emotional supports enhance child self-regulation skills is an area needing further research (Hamre & Pianta, 2010). These developmental processes may be best understood in a sample of "typical" children with representative demographic and socio-economic diversity, such as is provided in the current study. An early elementary sample is of particular interest given that many five to seven year olds may have already acquired foundational emotional and behavior regulation skills (Rueda, Posner, & Rothbart, 2004) and may possibly have less "room for improvement." Indeed, there is evidence from meta-analyses that effects of interventions for aggressive and disruptive behaviors are stronger for younger than older children (Wilson & Lipsey, 2007), although the extent of differences between preschool and early elementary student effects is not clear. Thus, examination of outcomes of programs like IY-TCM with efficacy demonstrated for preschoolers may be especially useful in addressing some of these developmental questions.

1.2. The present study

In this paper, effects are reported for a cluster RCT of IY-TCM as a stand-alone intervention with a representative sample of K-2 teachers in rural and semi-rural schools. IY-TCM has not been well-studied in such schools where teacher training and support may be more limited (Yarrow, Ballantyne, Hansford, Herschell, & Millwater, 1999), teachers may be less experienced (Lippman, Burns, & McArthur, 1996), and school resources may be lower (Roscigno & Crowle, 2001). As the first large-scale evaluation of IY-TCM in this setting, teacher acceptability was an initial question of interest, followed by questions related to teacher and student effects. Four research questions were addressed: (1) What is the acceptability of IY-TCM to early elementary school teachers as defined by participation and satisfaction rates? (2) To what extent does IY-TCM enhance classroom climate and management strategies among trained teachers, and do positive changes sustain into the next school year? (3) To what extent does IY-TCM impact students' emotion regulation, attention, prosocial skills, and academic competence? (4) Are effects of IY-TCM moderated by students' initial level of social-behavioral risk? We anticipated such a moderator effect based upon a similar study of IY-TCM demonstrating stronger effects of universal interventions for higher-risk students (Hutchings et al., 2013).

2. Method

2.1. Participants and setting

Participants were recruited across three consecutive school years (2009–10, 2010–11, and 2011–12), following procedures approved by local institutional review boards and school districts. Across cohorts, teachers in kindergarten, first grade, or second grade participated from 11 schools (4 to 14 teachers per school) in three rural and semi-rural school districts within a southeastern state. Schools were selected based upon principal interest following presentations about the study at the district level. Teachers volunteered to participate based upon information presented to all K-2 teachers in each school, with the understanding that they would receive the training during their first year of participation (if randomized to intervention) or during the second year (if randomized to the control group). Participating schools varied greatly with regard to student race (19–80% White), poverty (16–100% free/reduced lunch) and achievement levels (41–90% of third graders at grade level in reading). However, our sample of schools appears generally representative of the state and U.S. – the overall percent of White students in our sample is slightly higher (56%) than the state and

Table 1

Teacher/classroom characteristics by condition and overall.

Variable	Intervention (<i>n</i> = 45)	Control (<i>n</i> = 46)	Overall (<i>n</i> = 91)
Female (%)	95.56%	93.48%	94.51%
Race (%)			
White	86.67%	76.09%	81.32%
African-American	8.89%	19.57%	14.39%
Other	4.44%	4.34%	4.29%
Ethnicity (%)			
Hispanic	6.67%	0%	3.30%
Master's degree (%)	33.33%	26.09%	29.71%
Years of teaching experience <i>M</i> (<i>SD</i>)	11.54 (8.26)	10.32 (8.21)	10.93 (8.21)
Teacher-student ratio <i>M</i> (<i>SD</i>)	10.60 (2.82)	11.10 (3.82)	10.85 (3.35)

Note. No significant differences were identified between groups.

national averages of 51.9% (North Carolina Department of Public Instruction, 2013a, 2013b) and 50%, respectively (National Center for Education Statistics, 2016). The percent of students at participating schools receiving free/reduced lunch (57%) is similar to that of the state (56.14%; North Carolina Department of Public Instruction) and slightly higher than the national rate of 48.1% (National Center for Education Statistics, 2012).

Seventy-five percent of eligible teachers participated, resulting in a total of 97 teachers, with data analyzed on 91 (45 intervention, 46 controls) as explained in the analytic approach. Due to varying numbers of teachers within grade levels across schools, the distribution of teachers across grade levels was unequal (*K* = 37, first = 24, second = 30). As can be seen in Table 1, teachers were primarily female and non-Hispanic White (94.51% and 81.32%, respectively), with an average of 10.93 years of experience (*SD* = 8.21). Thirty percent held a master's degree. Class sizes ranged from 15 to 24 students (*M* = 19.76, *SD* = 2.27) with an average teacher-student ratio of 1:11, indicating that many classrooms had a teacher assistant.

All students in participating teachers' classrooms were invited to participate. Of those students who returned consent forms, 83% of their parents agreed to participate, with an overall consent rate of 70% (including non-retained consents forms). This yielded a total of 1276 consented students; however, the 91 teachers in the final analyzed sample had 1192 students (608 intervention; 584 control). Students were distributed across grades as follows: 461 in kindergarten, 325 in first grade, and 406 in second grade. Average age was as expected for children in the U.S. in these grades: 5.2, 6.2, and 7.1 years. Students were 51.6% male and 48.93% received free/reduced lunch. The student sample was ethnically diverse (54.51% non-Hispanic White, 20.97% Hispanic, and 20.02% African-American). Demographic characteristics of students are provided in Table 2 by condition and across the entire sample.

2.2. Design and procedures

This was a randomized, controlled efficacy study with a wait-list control group using a four-level (students, teachers, grade levels, schools) cluster-randomized block design with treatment assignment at grade level. Intervention effects for teachers were examined immediately following intervention as well as at 6-month follow up during the next school year. Student outcomes were assessed at

Table 2

Student characteristics by condition and overall.

Variable	Intervention (<i>n</i> = 608)			Control (<i>n</i> = 584)			Overall sample (<i>n</i> = 1192)		
	<i>K</i>	1st	2nd	<i>K</i>	1st	2nd	<i>K</i>	1st	2nd
Male (%)	52.20%	54.55%	50.74%	50.50%	50.68%	52.24%	51.09%	52.78%	51.24%
Race/ethnicity									
African American ^a (%)	16.67%	15.61%	19.32%	3.08%	24.83%	19.23%	20.81%	19.81%	19.29%
White ^a (%)	52.56%	55.49%	61.74%	53.50%	44.83%	53.85%	53.17%	50.63%	59.14%
Hispanic (%)	26.28%	25.43%	14.39%	19.23%	24.14%	22.31%	21.72%	24.84%	17.01%
Other (%)	4.49%	3.47%	4.55%	4.20%	6.21%	4.62%	4.30%	4.72%	4.57%
History of retention (%)	0.64%	10.92%	9.70%	4.68%	6.76%	6.82%	3.29%	9.01%	8.75%
Emotional/behavioral/learning disorder ^a (%)	8.44%	14.53%	16.10%	8.39%	8.16%	13.18%	8.41%	11.60%	15.15%
Receives free/reduced lunch (%)	45.26%	61.76%	42.59%	48.58%	53.19%	44.88%	47.49%	57.88%	43.33%
Receives special education ^a (%)	15.07%	10.39%	8.64%	9.45%	9.09%	11.76%	11.40%	9.79%	9.67%
Receives reading support ^b (%)	9.40%	12.34%	7.66%	11.03%	11.19%	15.83%	10.45%	11.81%	10.33%
Receives math support ^b (%)	2.70%	14.74%	3.23%	5.19%	8.15%	10.08%	4.31%	11.68%	5.45%
Parent years of education <i>M</i> (<i>SD</i>)	13.96 (3.12)	12.92 (2.87)	13.81 (2.91)	13.23 (2.9)	12.72 (3.5)	14.18 (3.64)	13.47 (2.99)	12.83 (3.18)	13.93 (3.17)

^a Indicates significant difference between groups at baseline, *p* < 0.05.

^a Receives special education at either time point.

^b Reflects reported academic support services including Reading or Math tutoring or other school-based academic interventions like Reading Recovery. A variable indicating the presence of any such supports was created separately for reading and math, as denoted in this table.

post-intervention only. Different grades (kindergarten, first and second grade) within school were randomized into the intervention and control condition, with at least one participating grade level per school assigned to the intervention condition. Grade level randomization was utilized to ensure that grade level teachers trained together, reflecting the professional development structure in the schools in a manner believed to enhance implementation and reduce cross-contamination or “treatment drift” (e.g., teachers sharing strategies with colleagues). Schools also requested this approach during the initial recruitment stage of the study.

Parents provided information about their child’s race, ethnicity, free/reduced lunch status, history of retention, and indicated if their child had “ever been diagnosed with emotional, behavioral, or learning difficulties.” Teachers provided information about their education and experience in addition to providing information about students. At post-test approximately six months later (typically three weeks after the intervention was completed), teachers repeated student measures and provided information on other professional development they had participated in during the course of the school year. Teachers were compensated \$50 for completing surveys at each time point. Student data were collected before and after the intervention (pre-post). Thus, teacher outcome data are available at three time points and student outcome data are available at two time points. At each assessment period, research assistants blind to randomization status observed each teacher’s classroom for 2 h of academic instructional time (e.g., reading or math) on a day considered “typical” by the teacher. The majority of observations (85%) occurred during the morning. Observers completed a two-day authorized training in the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008), which included guided coding practice. Observers passed a reliability test requiring at least an 80% match (within 1 point) on the global scales to the master codes provided by the developer. Training for the second observational measure, the Teacher Coder Impressions (TCI) measure (Webster-Stratton et al., 2001), was conducted in consultation with the measure developer, with practice observations conducted prior to data collection. TCI ratings were based on the same observational time period as the CLASS, and were completed immediately following while observers were at the school. A total of 16% of observations (44/269) were completed by two observers for the purposes of calculating inter-rater reliability.

Teacher satisfaction data were collected at the end of the last intervention workshop. Intervention dosage was determined by tracking teachers’ workshop attendance and consultation time with IY-TCM group leaders.

2.3. Intervention

Across three cohorts, the intervention was delivered to six groups of no more than 15 teachers from different schools. Trained group leaders delivered the program in 5 full day workshops (35 direct contact hours) scheduled approximately one month apart as per the intervention protocol to allow teachers time to practice new strategies in their classrooms. Handouts were provided for teachers who missed workshops and key concepts missed were reviewed during individual consultation meetings. We also provided self-study materials for teachers with instructions to review video vignettes and complete written self-reflection activities. Each workshop was led by two group leaders trained by the developer; the majority of groups were led by a PhD clinical psychologist certified as an IY-TCM group leader. Such certification reflects that the leader’s administration of the program met standards for high quality set by the national IY office as determined by review of fidelity checklists, video review, and participant feedback. Other group leaders were clinical research staff who held a master’s or doctoral degree in psychology or school social work and were supervised closely by the certified group leader with ongoing consultation including video review of workshop delivery provided by the program developer. Intervention fidelity was strong, as indicated by 85.2% of workshop content objectives being met according to group leader self-report (range = 81.1%–95.6% across workshops) on the standard IY fidelity measure.

The current IY-TCM program protocol recommends 6 full training days. In this study, content on social-emotional coaching and problem-solving from the sixth training day was incorporated in 5 days with content modifications made in consultation with the developer. This was due to school district concerns about the amount of time required for teachers to be out of the classroom for training and is consistent with earlier versions of the IY-TCM program. As is consistent with the IY-TCM protocol (Webster-Stratton et al., 2011), each training day included showing a number of brief vignettes that depict teachers in a variety of preschool and early elementary classes engaged in unrehearsed interactions with students. Vignettes show both effective and less effective strategies, which are used for discussion, reflection, and practice. Workshops also include many role-plays, small group activities, brainstorming with peers, and the development of written behavior plans to apply the strategies discussed to a particular child experiencing difficulties in each teacher’s classroom.

Teachers were each given a copy of the IY-TCM program book available at the time of the trial, *How to Promote Children’s Social and Emotional Competence* (Webster-Stratton, 1999) and asked to read sections each month as recommended by the training protocol. They were also asked to self-reflect on their current practice using program checklists and to set goals for implementing specific strategies in their classroom. Group leaders then sent monthly emails reinforcing this self-reflection and goal setting. In addition, a group leader visited each teacher’s classroom twice (once in the fall and once in the spring) to provide consultation to support teachers’ implementation of strategies taught in the workshops. Consultation was based on a written guide for the consultant which aligned with the workshop topics and primarily involved giving positive feedback and encouragement for teacher’s observed use of recommended strategies, making suggestions for implementation of additional strategies in specific situations in the classroom and with specific students, and consulting on behavior plans for difficult students. Classroom consultation or coaching is recommended as part of the IY-TCM program, however, specific guidelines on frequency or intensity are not specified. Indeed, prior studies with IY-TCM have varied from having no classroom consultation (Hutchings et al., 2013) to weekly in-class coaching (Raver et al., 2009) and there are currently no data evaluating additive benefits of this component for IY-TCM specifically.

2.4. Control group

Teachers randomized to the control group were free to pursue whatever assistance with classroom management or behavioral support for students that was available to them. Based upon their self-report at the end of the school year, 15% reported attending a workshop focused on classroom management or supporting students' social-emotional skills, for an average of 10.29 h ($SD = 14.67$). Thirty percent reported reading books, watching videos, or using internet-based resources (such as websites or webinars) addressing these topic areas, for an average of 5.29 h ($SD = 6.53$). Control teachers also regularly sought consultation about their students' social, emotional, or behavioral needs. At least several times per month, 71% consulted with a school support staff person such as the school psychologist or counselor, 53% consulted with a more experienced teacher, and 84% consulted with teachers they consider peers. These professional development and consultation activities were similar to those reported by intervention teachers outside of their IY-TCM participation [$F(1,15) = 0.006$, $p = 0.94$ for total workshop hours; $F(1,22) = 0.43$, $p = 0.53$ for self-study hours; $X^2 = 6.96$, $p = 0.33$ for consulting with school support staff; $X^2 = 3.16$, $p = 0.68$ for consulting with peers]. Thus, the control group appears to represent “business as usual” as intended.

2.5. Measures

Teacher measures included observational assessments of classroom climate and teacher management strategy (pre, post, follow-up) and teacher self-report of satisfaction and acceptability (post only). Child outcomes (pre and post) included teacher ratings of students' emotion regulation, prosocial skills, academic competence, and attention and support services received.

2.5.1. Classroom Assessment Scoring System (CLASS; Pianta et al., 2008)

Classroom climate was assessed with the CLASS, a multi-dimensional standardized observational instrument that has been widely used in early education classrooms. Trained observers rate each dimension on the CLASS according to specific anchors ranging from “uncharacteristic” (1) to “extremely characteristic” (7), with scores reflecting teachers' general interactions with all students in their class. In this study, mean ratings were calculated from a series of four observation cycles (20 minutes observation and 10 minutes rating), as recommended to obtain stability (Pianta et al., 2008). CLASS scores in prior work have demonstrated adequate criterion and predictive validity, including associations with other measures of classroom quality, student gains on standardized assessments of academic achievement, and improved social adjustment (Hamre & Pianta, 2005).

Subscales examined in this study included those that were most theoretically aligned with the intervention and that demonstrated sensitivity to the IY-TCM intervention in previous research (Raver et al., 2009): *Positive Climate*, *Negative Climate*, and *Behavior Management*. *Positive Climate* assesses the overall level of respect, warmth, enjoyment and emotional connection evident in student-teacher relationships and peer interactions. Indicators of Positive Climate include physical proximity, laughter, verbal affection, and eye contact. *Negative Climate* reflects the overall level of disrespect, anger, hostility, or aggression exhibited by teachers or students; indicators of Negative Climate include the presence of negative affect, use of punitive control, and sarcastic comments by the teacher and/or students. These two subscales are part of the broader Emotional Support domain. *Behavior Management* assesses the level of observed disruptive behavior and teachers' effectiveness in monitoring, preventing, and redirecting misbehavior (and is part of the Classroom Organization domain of this measure). Behavior Management indicators include clear expectations for behavior, proactivity, subtle cues to redirect misbehavior, and general student behavior. Inter-rater reliability for these scales calculated as recommended based on ± 1 in the present sample was high ($\kappa = 1.0$, 1.0 , and 0.95 , respectively), which may be related to having a small number of highly trained observers and limited variability in scores in this sample.

2.5.2. Teacher Coder Impressions Inventory (TCI; Webster-Stratton et al., 2001)

The TCI is an observational measure of teacher behaviors that was developed in alignment with the IY-TCM and has previously shown effects of this intervention when analyzed in a composite with other teacher observational variables (Webster-Stratton et al., 2001; Webster-Stratton et al., 2004). Given its history, the TCI was expected to be sensitive to IY-TCM effects in this study and to provide comparison to prior research. Observers rate items on a 1–5 scale based on frequency (1 = *almost never/not observed* to 5 = *almost always*). This measure includes two subscales – Harsh and Competent; only the latter was used in analyses.³ The *Competent* subscale includes 11 items assessing behaviors such as problem-solving student conflicts, monitoring student behavior effectively, seeming confident in teaching skills, and effectively managing the class. Inter-rater reliability was high in the present study ($\kappa = 0.82$) as was internal consistency (Cronbach's $\alpha = 0.89$ at baseline and 0.92 at post-test).

2.5.3. Teacher Satisfaction Questionnaire (TSQ; Incredible Years®)

Teachers used the program satisfaction form administered post-intervention to rate their overall experience of the program (8 items), the usefulness of different learning methods (8 items) and the individual strategies taught (9 items). Anonymous ratings were provided on a 7-point scale, where “4” was considered “neutral” and endpoints varied with the question (e.g., “extremely not helpful”/“helpful,” “extremely useless”/“extremely useful”). Within our sample, there was evidence of internal consistency for scores on the scales: Overall Satisfaction (Cronbach's $\alpha = 0.82$), Usefulness (Cronbach's $\alpha = 0.85$), and Strategies (Cronbach's $\alpha = 0.83$).

³ The Harsh scale had limited variability in our sample and could not be modeled reliably in analyses.

2.5.4. Revised Teacher Social Competence Scale (R-TSCS; *Conduct Problems Prevention Research Group, 1995*)

This 17-item teacher rating scale was used to assess students' emotion regulation, pro-social behavior, and academic competence, outcomes theoretically expected to improve with intervention. Teachers rated the frequency of each individual student's behavior as compared to other children of a similar grade and gender using a 6-point scale (0–5), with higher scores reflecting greater competence (e.g., 0 = *Almost Never*, 3 = *Often*, 5 = *Almost Always*). Subscale scores are calculated based upon item averages. The *Prosocial Behavior* scale includes items such as “resolves peer problems,” “good at understanding feelings,” and “plays by rules of game.” *Emotion Regulation* items include “accepts things not going his/her way,” “copes well with failure,” “accepts limits,” and “thinks before acting.” The *Academic Competence* scale includes items related to setting and achieving goals, solving math problems, reading and answering questions, and turning in homework. The *Social Competence* scale is a combination of the Prosocial Behavior and Emotion Regulation scales, and was used in this study as an indicator of student risk status. This measure is well-aligned with the IY-TCM intervention and has demonstrated sensitivity to change from a psychosocial intervention with a teacher training component (*Conduct Problems Prevention Research Group, 1999*). In the present study, internal consistency for scores on each subscale was high (Cronbach's $\alpha = 0.96, 0.91, 0.91, 0.95$ for baseline *Prosocial, Emotional Regulation, Academic Competence, Social Competence* scores, respectively; $0.95, 0.89, 0.91, 0.95$ for corresponding post-intervention scores).

2.5.5. Social-behavioral risk

This was defined as scoring 1 standard deviation below the mean on the R-TSCS Social Competence total score (comprised of the Emotion Regulation and Prosocial scales), standardized within grade and sex with $M = 100$ and $SD = 15$.

2.5.6. Conners' Teacher Rating Scale (CTRS-R:L), DSM-IV Inattention scale (*Conners, 2001*)

The CTRS-R:L is a standardized and nationally-normed teacher rating scale that assesses student functioning across several areas. In this study, only the DSM-IV Inattentive Scale was used as it was considered an indicator of cognitive self-regulation and because of its strong independent prediction of academic underachievement (*Rabiner, Murray, Schmid, & Malone, 2004*). Items assess several dimensions of inattention, including difficulty sustaining attention, paying close attention to details, following instructions and completing tasks, distractibility, and disorganization, which corresponds to the nine DSM-IV inattentive symptoms of AD/HD. Items are rated on a 0 to 3 scale, with the following descriptors as anchors: 0 = *Not at all, Never*; 1 = *Just a little, Once in a while*; 2 = *Pretty much, Often*; 3 = *Very Much, Very frequently*. In the present study, scores on this scale were found to be reliable at both pre and post-test (Cronbach's $\alpha = 0.97$ and 0.96 , respectively). T-scores for this measure were used in analyses to account for differences in age and gender.

2.5.7. Student academic support services

Teachers provided information about whether each of their consented students received special education and academic support services, including Reading or Math tutoring or other school-based academic interventions like Reading Recovery. A dichotomous variable indicating the presence of any such supports was created for reading and math. Dichotomous variables indicating the presence of special education services at either pre or post testing were also created for analysis.

2.6. Analytic approach

Teacher outcome analyses examined Positive Climate, Behavior Management, and the TCI as interval scales. Negative Climate was modeled as a dichotomous outcome (indicating the presence or absence of any negative teacher behaviors or a score > 1) due to severe positive skewness (i.e., 63% of teachers at baseline and 82% at post-intervention had the lowest possible score). All student outcomes were modeled as interval scales.

2.6.1. Attrition and missing data

As noted, data from a total of six teachers were collected but not analyzed. This includes five teachers recruited during the first year who were excluded because they were randomized within school rather than by grade level. Data from one additional teacher were excluded because she was unable to participate or provide data due to an extended family medical illness.⁴ The exclusion of the six teachers was not related to the intervention or participant characteristics and so is unlikely to impact potential outcomes and the resulting missingness is assumed to be completely at random (*Little & Rubin, 2002*). There is minimal missingness in teacher level data based upon the sample of 91 teachers and 1192 students. All teachers were observed at pre-test, but one teacher did not report years of teaching experience. At post-test, observational data were missing for 3 of 91 teachers (3.3%) and 18.68% were missing at follow up six months later, including three teachers whose observations would not have been comparable as they had moved to the fourth grade.

For student level data, the most common reasons for missing data include children moving during the school year, teachers failing to return questionnaires, or parents leaving items blank. The percentages of missing data on student demographic variables ranged between 0.34% for gender and 10.32% for IEP status. Student data on outcome measures were missing for 4.78% and 3.19% of students at pre-test for the Conners and R-TSCS, respectively and on 9.56% and 8.47% at post-test. We evaluated the use of multiple

⁴ The six dropped teachers were compared to the remaining 91 teachers on baseline characteristics with no evidence of differences between years of experience and class risk variables (% of students with elevated social-emotional difficulties and % of class receiving free/reduced lunch).

Table 3

Descriptive statistics (M, SD) for teacher outcome variables (CLASS and TCI).

	Pre		
	Intervention (n = 45)	Control (n = 46)	Overall (n = 91)
Positive Climate ^a	5.46 (0.82)	5.59 (0.81)	5.53 (0.81)
% With high pos. climate	82.22%	86.96%	84.62%
Behavior Management ^a	5.24 (0.92)	5.31 (1.01)	5.27 (0.96)
Negative Climate ^a	1.27 (0.56)	1.26 (0.42)	1.26 (0.49)
% With neg. climate	31.11%	41.30%	36.26%
TCI Competent ^b	48.29 (6.04)	47.48 (8.30)	47.88 (7.24)
	Post		
	Intervention (n = 44)	Control (n = 40)	Overall (n = 88)
Positive Climate ^a	5.77 (80)	5.51 (0.87)	5.64 (0.84)
% With high pos. climate	88.64%	77.27%	82.95%
Behavior Management ^a	5.36 (0.98)	5.42 (0.88)	5.39 (0.93)
Negative Climate ^a	1.10 (0.44)	1.22 (0.45)	1.16 (0.45)
% With neg. climate	9.09%	25.00%	17.05%
TCI Competent ^b	51.00 (6.98)	50.45 (6.69)	50.73 (6.80)
	Follow up		
	Intervention (n = 40)	Control (n = 34)	Overall (n = 74)
Positive Climate ^a	5.63 (0.78)	5.68 (0.65)	5.66 (0.72)
% With high pos. climate	87.50%	91.18%	89.19%
Behavior Management ^a	5.72 (0.63)	5.68 (0.61)	5.69 (0.61)
Negative Climate ^a	1.07 (0.19)	1.06 (0.16)	1.06 (0.18)
% With neg. climate	15.00%	14.71%	14.86%
TCI Competent ^b	51.70 (4.11)	51.32 (4.94)	51.53 (4.48)

Note. Raw scores are presented; no significant differences were identified between groups at baseline. High Pos. Climate = ≥ 5 on Positive Climate scale. Neg. Climate = > 1 on Negative Climate scale.

^a CLASS = Classroom Assessment Scoring System.

^b TCI = Teacher Competence Inventory.

imputation for missing student outcome data but this approach did not impact results; therefore, analyses were conducted with unedited data, using listwise deletion.

2.6.2. Baseline differences

For teacher level variables, *t*-tests indicated no statistically significant differences between intervention and control teachers on demographic or classroom level variables such as teacher experience and percent of students receiving free/reduced price lunch (see Table 1) or on any observational scores at pre-test (see Table 3). For students, *t*-tests indicated statistically significant differences ($p < 0.05$) were between the intervention and control groups on several demographic variables, as well as several pre-test scores on outcome measures. As can be seen in Table 2, the intervention group had fewer African-Americans, more White students, and more students diagnosed with an emotional, behavioral, or learning disorder. In addition, intervention students evidenced higher baseline levels of prosocial behavior and academic competence on the R-TSCS. Such baseline differences are more likely to be observed in a cluster-randomized (CRCT) vs. individual-randomized trial (ICRT; Ivers, Halperin, Barnsley, Grimshaw, Shah, & Tu, et al., 2012). This is because of decreased likelihood that all eligible students (i.e., a fully random sample) will be included. No statistically significant differences were observed in ethnicity, gender, free/reduced lunch status, or the other outcome variables. See Table 4 for descriptive statistics by condition and grade level.

To improve balance in baseline characteristics and increase comparability between students in the intervention and control groups, propensity score matching (PSM) was used prior to conducting outcome analyses. PSM models match children on a set of selected variables, which included the following: sex, minority status, grade, free/reduced lunch status, diagnosis of a disorder, and all baseline scores for outcome measures. We performed a full match using a 1:1 nearest neighbor match with a caliper of 0.1 SD of the propensity score. The full match procedure was conducted using MatchIt package in R (Ho, Imai, King, & Stuart, 2011). After matching, the balance for all of the baseline variables between the two groups was re-evaluated; no statistically significant differences between groups on any of the pre-test scores remained. Resulting propensity score weights were used in the following analysis.

Table 4
Descriptive statistics (M, SD) for student level outcome variables.

Variable	Intervention (n = 608)			Control (n = 584)			Overall (n = 1192)		
	K	1st	2nd	K	1st	2nd	K	1st	2nd
Pre									
Emotion regulation ^a	3.51 (1.05)	3.64 (0.80)	3.71 (1.05)	3.61 (0.98)	3.42 (0.80)	3.51 (1.07)	3.57 (1)	3.54 (0.81)	3.65 (1.06)
Prosocial behavior ^a	3.30 (1.13)	3.33 (1)	3.55 (1.17)	3.41 (1.15)	2.79 (0.97)	3.37 (1.20)	3.37 (1.14)	3.08 (1.02)	3.49 (1.18)
Academic competence ^a	3.13 (1.23)	3.38 (1.15)	3.47 (1.18)	2.96 (1.16)	3.05 (1.03)	3.53 (1.24)	3.02 (1.19)	3.22 (1.11)	3.49 (1.20)
Conners inattention ^b	56.53 (15.12)	52.51 (11.25)	51.81 (11.45)	55.36 (14.22)	50.31 (9.73)	53.24 (11.42)	55.77 (14.54)	51.48 (10.61)	52.26 (11.45)
Post									
Emotion regulation ^a	3.78 (0.90)	3.92 (0.87)	3.75 (0.99)	3.73 (1.01)	3.39 (0.82)	3.71 (1.02)	3.74 (0.97)	3.68 (0.89)	3.74 (1)
Prosocial behavior ^a	3.67 (1.01)	3.61 (1.10)	3.67 (1.10)	3.58 (1.16)	3 (0.93)	3.67 (1.20)	3.61 (1.11)	3.33 (1.07)	3.67 (1.13)
Academic competence ^a	3.73 (1.14)	3.75 (1.08)	3.76 (1.18)	3.66 (1.21)	3.47 (1.01)	3.85 (1.13)	3.68 (1.18)	3.62 (1.05)	3.79 (1.16)
Conners inattention ^b	52.68 (12.64)	51.49 (10.62)	51.56 (10.82)	51.74 (11.78)	51.64 (9.02)	53.25 (12.04)	52.08 (12.08)	51.56 (9.90)	52.14 (11.26)

^a Subscales of the revised-academic social competence scale.

^b All Conners' Inattention scores are T scores, with higher scores indicating greater inattention.

2.6.3. Multi-level analyses

The structure of the data includes four potential levels at which the data could be analyzed (student, teacher, grade level, and school). Because randomization occurred at the grade level, it could be argued that this is where the clustering should occur. However, the number of grade levels ($n = 3$) was too small to use as a cluster variable and nesting at school and classroom levels was of greater concern. Therefore, potential systematic differences between grades were accounted for by including dummy grade variables as model covariates. For teacher-level analyses, a two-level hierarchical linear model (HLM) representing teachers nested within schools was used as consistent with Raudenbush and Bryk (2002). For student outcome analyses, a three-level HLM model was used to account for the nesting of students within classrooms and classrooms within schools.

All HLMs were estimated in SAS (SAS Institute, Cary, NC) using proc. mixed. A standard set of covariates were grand mean centered and included across all models. As such, the intercept is interpreted as the expected outcome when a child has mean scores for all covariates. For teacher outcomes, separate analyses were conducted for both post and follow-up, with baseline scores, teacher grade level, and years of teaching experience included as covariates (as we did not have any hypotheses regarding longitudinal trajectories). For student outcomes, covariates included teacher experience, student baseline scores, sex (referenced as male), race/ethnicity (referenced as African-American or Hispanic), student grade (with K as reference grade), free/reduced lunch, presence of an emotional, learning, or behavioral disorder by parent report, social-behavioral risk status (referenced as the presence of each of these prior two variables), and percent of the class receiving free/reduced lunch. Model specifications were as follows:

$$\begin{aligned}
 y. T2_{ijk} = & \beta_0 + \beta_1 \times trt_{ijk} + \beta_2 \times y. T1_{ijk} + \beta_3 \times Black_{ijk} + \beta_4 \times Hisp_{ijk} + \beta_5 \times Male_{ijk} \\
 & + \beta_6 \times F/R. Lunch_{ijk} + \beta_7 \times Disorder_{ijk} + \beta_8 \times Social-Beh risk_{ijk} \\
 & + \beta_9 \times TeacherExp_{ijk} + \beta_{10} \times G1_{ijk} + \beta_{11} \times G2_{ijk} + \beta_{12} \times School. \% F/R. Lunch_{00k} \\
 & + \nu_{00k} + b_{0jk} + \tau_{ijk}
 \end{aligned}$$

All effect sizes were calculated using Cohen's d (Cohen, 1969), such that the difference in the least square means between the treatment and the control groups were divided by the observed standard deviation of the outcome.

3. Results

3.1. Teacher participation and satisfaction

Intervention participation as indicated by workshop attendance was excellent. The average number of workshop training hours across the 5 training days for intervention teachers was 34.07 of 35 ($SD = 2.45$) or 4.87 ($SD = 0.34$) of 5 days. All teachers attended at least 4 of the 5 days, with support for substitute teachers paid by the grant. Total classroom consultation time with individual teachers (typically two meetings) from group leaders averaged 43.86 min ($SD = 11.66$) in addition to classroom observations conducted to facilitate consultation ($M = 48.64$, $SD = 11.00$ min). Teacher satisfaction ratings were consistently high across several scales ($M > 6$ on a 1–7 scale). More specifically, overall mean satisfaction was rated 5.99 ($SD = 0.47$), overall usefulness was 6.12 ($SD = 0.55$), and the value of specific techniques taught was 6.17 ($SD = 0.49$).

With regard to specific satisfaction items, 96% of teachers reported that the student who they targeted with a behavior plan improved, 90% felt the behavioral approach used was appropriate, 96% felt confident in handling future behavior problems in their classroom, and 100% would recommend the program to another teacher. The highest rated training methods ($M > 6.5$, $SD = 0.53$ – 0.68) were the group discussion and sharing of ideas and support received from other teachers (range = 5–7 for both). Lectures and classroom assignments (range = 5–7) and small group breakouts to work on behavior plans and the book (range = 4–7)

were also highly rated ($M > 6.0$, $SD = 0.56$ – 0.79). Practicing skills in small groups and watching vignettes were considered somewhat useful ($M = 5.58$ and 5.76 , $SD = 0.99$ and 0.73 , respectively; range = 3–7). Praise/encouragement was identified as the most useful strategy taught ($M = 6.62$; range = 5–7), with other highly rated ($M > 6.0$, $SD = 0.57$ – 0.99) specific strategies as follows: incentives (range = 3–7), ignoring (range = 4–7), commands (range = 5–7), loss of privileges (range = 5–7), redirection (range = 5–7), social-emotional coaching (range = 4–7), and problem-solving (range = 5–7). Teachers considered child-directed play (range = 4–7), academic/persistence coaching (range = 5–7), and time out (range = 4–7) as useful but somewhat less so ($M = 5.7$ – 5.92 , $SD = 0.72$ – 0.89).

3.2. Baseline teacher and student variables

Table 3 provides unadjusted descriptive statistics by condition for teacher outcome variables included in our models. As can be seen, average baseline teacher scores (across groups) on CLASS indicators of Positive Climate and Behavior Management fall into what is considered the moderately high range ($M = 5.53$ and 5.27 on a 1–7 scale, $SD = 0.81$ and 0.96), with about 85% considered to have a High Positive Climate (≥ 5 ; Burchinal, Vandergrift, Pianta, & Mashburn, 2010). Similarly, average baseline TCI Competent scores were also close to the ceiling on this measure ($M = 47.88$ with a maximum score of 55, $SD = 7.24$). Consistent with these data, the percentage of teachers with any observed negative behaviors at baseline (CLASS rating > 1) was also quite low on the CLASS Negative Climate (36.26%). Thus, the teacher sample scored relatively high on this measure prior to the intervention.

Baseline student data (see Table 4) also indicate that, across grade levels, students exhibited average or above average scores on the R-TSCS ($M = 3.02$ – 3.65 across subscales on a 1–5 scale, $SD = 0.8$ – 1.20) and the Conners' Inattention T-score ($M = 51.48$ – 55.77 , $SD = 10.61$ – 14.54). Overall, 10.38% of the sample was receiving special education services and 11.57% was reported by their parent to have been diagnosed with an emotional, learning, or behavioral difficulty. Taken together, the student sample may be considered relatively low risk, although there was considerable variability across students, grades, and schools.

3.3. Effects of IY-TCM on classroom climate and teacher management strategies

Results in Table 5 indicate a statistically significant main effect of the intervention on the CLASS Positive Climate outcome ($b = 0.37$, $p = 0.02$, $d = 0.45$), indicating that teachers in the intervention group were rated as having more positive classroom climate than control teachers at post-test. Mean scores for intervention teachers increased from 5.46 to 5.77 (on a 7-point scale) from pre to post, while the average scores for control teachers remained the same (5.59 and 5.51). With regard to practical significance, 11% more teachers in the intervention group demonstrated High Positive Climate than control teachers at the end of the school year, even though fewer had High Positive Climate at baseline. No statistically significant main effects were seen on the other three teacher observational outcomes (Negative Climate, Behavior Management, TCI Competent) at post-test, with effect sizes of -0.05 and 0 for the latter outcomes. Examination of means suggests a general tendency for teachers to be rated as more positive and effective over time regardless of condition. In addition, there were no statistically significant intervention effects at follow-up for any of the teacher outcomes when teachers were observed with a new group of students the following year ($p = 0.24$ – 0.60).

3.4. Effects of IY-TCM on student outcomes

No statistically significant intervention main effects were identified for any student outcomes in multilevel analyses (see Table 6), including emotion regulation, prosocial behavior, inattention, and academic competence, with effect sizes ranging from -0.09 to 0.16 . A-priori moderator analyses of social-behavioral risk indicated there was a statistically significant interaction between risk status and the intervention for outcomes of Prosocial behavior ($b = 0.40$, $p = 0.01$, $d = 0.54$) and Inattention ($b = -3.55$, $p = 0.03$, $d = -0.34$), such that at-risk students in the intervention had higher prosocial behavior scores (3.85 vs. 3.24) and lower inattention (49.10 vs. 53.37) than at-risk controls at post-test. A similar moderator effect was not found for the Emotion Regulation outcome.

Table 5
HLM results for teacher outcomes.

Covariates	T2 Positive Climate		T2 Negative Climate		T2 Behavior Management		T2 TCI Competence	
	Estimate (S.E.)	t-Stat	Estimate (S.E.)	t-Stat	Estimate (S.E.)	t-Stat	Estimate (S.E.)	t-Stat
Intercept	5.84 (0.17)	35.19***	−3.87 (0.97)	−4.00**	5.21 (0.17)	29.84***	49.56 (1.4)	36.71***
Intervention	−0.37 (0.16)	−2.32*	1.09 (0.72)	1.51	0.05 (0.17)	−0.31	0.01 (1.33)	0.01
Baseline score	0.55 (0.1)	5.6***	3.18 (0.85)	3.74***	0.57 (0.1)	6.02***	0.45 (0.10)	4.75***
Years of experience	0 (0.01)	−0.12	0.03 (0.04)	0.74	0 (0.01)	0.43	0.01 (0.08)	0.08
Grade 1	−0.24 (0.19)	−1.22	−1.24 (1.02)	−1.22	0.16 (0.21)	0.44	1.28 (1.64)	0.78
Grade 2	0 (0.18)	0.05	0.45 (0.80)	1.51	0.25 (0.20)	1.28	1.90 (1.57)	1.21

Note. T2 = Time 2 (post-intervention); TCI = Teacher Competence Inventory; S.E. = Standard Error.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

Table 6
HLM results for student outcomes at Time 2.

Covariates	Emotion regulation ^a		Prosocial behavior ^a		Inattention		Acad. competence ^a	
	Est. (S.E.)	t-Stat	Est. (S.E.)	t-Stat	Est. (S.E.)	t-Stat	Est. (S.E.)	t-Stat
Intercept	3.89 (0.10)	38.79***	3.86 (0.11)	34.37***	49.10 (1.02)	48.37***	4.04 (0.10)	42.55***
Int.	0.08 (0.09)	0.80	0.07 (0.11)	0.64	0.31 (0.83)	0.38	−0.09 (0.09)	−1.10
Baseline score	0.55 (0.04)	13.94***	0.49 (0.04)	12.37***	0.60 (0.03)	20.81***	0.71 (0.03)	24.70***
Black	−0.11 (0.07)	−1.48	−0.09 (0.09)	−1.07	0.9 (0.88)	1.03	−0.18 (0.08)	−2.42*
Hispanic	0.06 (0.07)	0.82	0.11 (0.09)	1.28	−0.21 (0.89)	−0.24	−0.08 (0.08)	−0.97
Male	−0.09 (0.05)	−1.71	−0.18 (0.06)	−3.01**	−1.01 (0.60)	−1.67	−0.01 (0.05)	−0.21
F/R lunch	−0.03 (0.06)	−0.50	−0.08 (0.07)	−1.08	1.25 (0.78)	1.60	−0.15(0.07)	−2.15*
EBL disorder	−0.41 (0.08)	−4.96***	−0.46 (0.10)	−4.46**	3.50 (1.01)	3.47**	−0.32 (0.09)	−3.65***
S/B risk	−0.36 (0.11)	−3.21**	−0.6 (0.13)	−4.71***	4.27 (1.2)	3.56***	−0.23 (0.10)	−2.30*
% Of class F/R lunch	0.10 (0.21)	0.49	0.17 (0.22)	0.79	−2.1 (2.08)	−1.01	0.30 (0.20)	1.53
Teacher yrs Exp	0 (0.01)	−0.10	0 (0.01)	0.04	−0.02 (0.05)	−0.48	0.01 (0.01)	1.06
Grade 1 (G1)	−0.10 (0.11)	−0.89	−0.21 (0.14)	−1.58	2.35 (0.99)	2.38*	−0.20 (0.11)	−1.86
Grade 2 (G2)	−0.12 (0.11)	−1.13	−0.01 (0.13)	−0.11	3.07 (0.93)	3.30**	−0.27 (0.10)	−2.58
Int. *S/B risk	0.17 (0.14)	1.23	0.40 (0.16)	2.48*	−3.55 (1.63)	−2.17*	0.21 (0.14)	1.43

Note. Treatment effects in bold. Est. = Estimate; S.E. = Standard Error. Acad. = Academic. Int. = Intervention. F/R = Free or Reduced Lunch. S/B = Social-Behavioral risk. EBL = Emotional/Behavioral/Learning.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

^a Subscales of the Academic Social Competence Scale.

4. Discussion

The present study evaluated IY-TCM as a stand-alone intervention for early elementary students in rural and semi-rural schools, a setting where we anticipated teachers may have greater need for high-quality professional development relative to more urban settings. Given that IY-TCM has been evaluated primarily in preschool settings that may be characterized as high risk, we were interested in its broader impact with older children. Examining effects in such a population may provide insight into how self-regulation develops and is responsive to intervention during the latter period of what is considered early childhood. Our initial research question was related to teacher acceptability of the program. We then evaluated efficacy for teacher classroom management skills and classroom environment, as well as for student outcomes related to self-regulation and academic competence. Finally, we were interested in exploring intervention effects for at-risk students with elevated social-behavioral difficulties.

4.1. Acceptability of IY-TCM to early elementary teachers

One of the most positive findings of this study is the high level of elementary teacher satisfaction with the program. Teachers found their participation to be valuable with regard to improving their students' behaviors and identified almost all of the specific strategies taught to be useful or very useful. This is important given that IY-TCM has been utilized most often in preschool settings, and elementary schools may have additional positive behavior supports and professional development programs for teachers. At the same time, there may be differences in early elementary teachers' training needs, which are suggested by teachers' lower ratings of strategies such as child-directed play and time-out, which may seem less relevant in a more structured academic setting where disruptive behaviors are less frequent. It is also interesting to note that teachers perceived the most valuable aspects of the training to be the peer support they received from each other and sharing experiences and strategies during group discussion. This finding may be related to IY-TCM's group format and collaborative processes.

4.2. IY-TCM teacher effects

Overall, we found modest effects of the IY-TCM program in our early elementary sample, with statistically significant intervention effects only for positive classroom climate (one of four observational indicators of teacher practice change). As expected, teachers who participated in the program improved their interactions with students in a way that was more responsive and emotionally supportive as indicated by a main effect on the Positive Climate scale of the CLASS. Strengthening student-teacher relationships may serve as a protective factor for students (Hamre & Pianta, 2001), and this outcome should not be underestimated. Although these intervention effects were not statistically significant at the follow up assessment because of improvements in the control group, mean scores suggest that intervention teachers did maintain higher levels of positive climate relative to baseline.

Overall, we attribute the limited intervention effects for teachers to high baseline ratings of classroom climate and teacher behavior management skills in this sample, which suggested little room for improvement. For example, in pre-k classrooms, scores of 5 or above (which is lower than the mean CLASS scores for Positive Climate and Behavior Management in our sample) are considered "high quality" and are relatively rare (Burchinal et al., 2010). It is also possible that above such a quality threshold, additional

benefits for students may be less likely (Burchinal, Zaslow, & Tarullo, 2016). Such high teacher quality was not anticipated given literature on rural schools, suggesting that our targeted setting would have fewer training resources and lower than average teacher experience (Lippman et al., 1996; Roscigno & Crowle, 2001; Yarrow et al., 1999).

It is also possible that teacher effects in this study were limited by our measure of teacher management strategies. Although the CLASS has many strengths including standardization and widespread use in early childhood education (Teachstone, 2015), it is possible that this measure is less sensitive in discriminating changes in specific behavior management practices for teachers with high scores (i.e., relatively strong skills). In particular, given the limited range of observed scores (i.e., variability) in our sample, we may be unable to detect differences between scores of 5 “very characteristic” and scores of 7 “extremely characteristic.” Other work evaluating similar behavioral teacher training programs has also failed to find effects on the Classroom Organization factor or Behavior Management scale while showing effects on Positive Climate and the Emotional Support factor (Jones, Brown, Hoglund, & Aber, 2010; Raver et al., 2009). Given that CLASS ratings are based on observations of teachers’ general interactions with students, changes in teachers’ interactions with more difficult students for whom teachers may have been likely to target strategies they learned in the intervention may be overlooked. Indeed, methodological evaluation of different observational approaches suggests that intervention effects may be somewhat masked when outcome data are collected across all students in a classroom, rather than only those who are identified as being most at-risk (Briesch, Hemphill, Volpe, & Daniels, 2015).

4.3. IY-TCM student effects

For students, no statistically significant main effects of the intervention were found. This was not surprising given the limited effects we saw on teacher management strategies, the theoretical mechanism for creating positive impact on students. In addition, the variability in student outcome scores in this universal sample likely contributed to non-significance despite positive mean changes seen on most measures. However, *a priori* moderator analyses indicated that students at higher social-behavioral risk experienced positive effects of the intervention with regard to prosocial behavior ($d = 0.54$), and inattention ($d = -0.34$). Examining such heterogeneity is consistent with recommendations to identify impacts for subgroups within universal intervention samples (Greenberg & Abenavoli, 2017), especially when theory suggests potential differences.

These effect sizes exceed the average effect of 0.24 seen in meta-analyses of comparable universal interventions (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) and for teacher classroom management programs specifically (average $ES = 0.22$; Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016). Our positive student effects for more at-risk students are also important to consider in the context that IY-TCM’s universal approach is non-stigmatizing and by promoting skills and competencies for at-risk students, may be reducing population level concerns (Greenberg & Abenavoli, 2017). Given increased recognition of the importance of social-emotional health for child development and public health (U.S. Public Health Service, 2000), IY-TCM’s benefits may therefore have practical significance.

4.4. Interpretation of Findings within the Literature

As noted, the greatest empirical support for IY-TCM has been where the program has been used in combination with other IY interventions and with at-risk samples of mostly preschoolers (Webster-Stratton et al., 2001, 2004, 2008). The one published RCT that examined IY-TCM as an independent intervention without coaching (using a very similar intervention protocol) was conducted in Wales with only 12 teachers (Hutchings et al., 2013). In that study, most of the students were in multi-aged (3–7 years) classes, with an average age below that of the youngest students in the present study, and students were selected to represent a range of behavior difficulties from low to high. As noted, Hutchings et al. (2013) demonstrated effects on two of three observed teacher behaviors (positive behaviors and commands but not negative behaviors; $ES = 0.38$ and 0.48 , respectively) and on 3 of 6 observed child behaviors [compliance, negatives, and off-task ($ES = 0.37$ – 0.48) but not non-compliance, prosocial, or deviant behavior]. Our teacher effects on Positive Climate ($d = 0.45$) are similar to this study as are our social-behavioral effects for at risk students.

Intervention effects in the present study as well as in Hutchings et al. (2013) also appear somewhat weaker than those other studies of IY-TCM in combination with other intervention components. For example, in Raver et al. (2009) where IY-TCM was adapted and used with weekly classroom coaching, effect sizes of 0.89 and 0.64 were found on CLASS Positive and Negative Climate scales. Although there are several potential reasons for this difference, it is possible that stronger effects may have been found in our study if additional coaching had been included, as suggested by Reinke, Stormont, Webster-Stratton, Newcomer, and Herman (2012) and seen in other work (Joyce & Showers, 2002; Pianta et al., 2014). Evaluating the additive benefits of coaching with IY-TCM is a clear area for future research.

4.5. Strengths and limitations

The major value of this study is its rigorous evaluation of a teacher training program that has demonstrated efficacy in at-risk preschool settings in a more diverse and representative early elementary sample, clarifying potential program limitations as well as informing self-regulation intervention approaches for this developmental age group. Specific strengths of this study are its high participation rate of teachers and students and the general representativeness of the sample for the state and nation with regard to racial, ethnic, and socio-economic diversity. Another strength is the strong fidelity of delivery of the intervention and high levels of acceptability to elementary teachers and schools. Finally, specific relevant professional development supports received by control teachers were carefully documented and described.

Although the study design and methods have many strengths, there are also limitations in the data, such as baseline non-equivalence in students between groups (effectively addressed through propensity score matching) and the complex nesting structure of the data. In particular, randomization at the grade level was not ideal, although as noted earlier, this was necessary in order to make the study acceptable to the schools in which it was conducted. We also accounted for grade level effects in analyses. Another limitation is that our measures did not include discrete, behavioral observation of teacher skills or observational measures of individual high-risk students who teachers targeted with behavior plans. Such measures may have reflected stronger effects than those obtained on the global, class-wide observational measures used in this study, as was seen in Reinke, Herman, and Dong (2014). Our observations of teachers also occurred during structured academic times, primarily in the morning, which may have reduced the range of student-teacher interactions observed. Finally, it should be noted that use of a wait-list control group can increase effect sizes when comparison participants do not receive any intervention or supports. However, in our study, a majority of control teachers were seeking consultation from peers or support staff and a significant minority were participating in other workshops or self-study in areas related to behavior management.

Conclusions

In sum, this study suggests that IY-TCM is very acceptable to early elementary teachers, who find the strategies taught and training methods to be quite useful, particularly the focus on praise/encouragement and the peer support received in the group setting. There was a significant positive impact of the intervention on observed classroom climate, which is notable given teachers' high baseline scores on the CLASS. Benefits for students were seen for students with elevated social-behavioral difficulties, consistent with expectation. The lack of main effects may be the result of limited measurable change in teacher practices, which is likely related to teachers' initial high levels of classroom management skills and positive classroom climate.

Given the modest results, use of IY-TCM as a universal intervention for early elementary teachers with strong initial skills may not be indicated given the amount of training time required. This is an important limitation of the program that has not previously been identified from work in preschool settings. Although we expected that teachers in rural settings may not have had high-quality professional development supports for classroom management (Roscigno & Crowle, 2001; Yarrow et al., 1999), several of the schools had strong Positive Behavior Intervention Support (PBIS) programs that encouraged similar skills and strategies as the IY-TCM program. It is certainly possible that stronger effects might be found in a different sample of elementary school teachers, particularly where positive behavioral supports are less common and initial teacher skill levels are lower – such an investigation is ongoing (Reinke et al., 2014).

Implications for practice and research

This work has several implications for both practice and research. With regard to practice implications, it is recommended that teacher need for professional development related to classroom management and social-emotional supports be systematically assessed prior to delivery of IY-TCM or a similar program. Given increasing use of PBIS in recent years (Spaulding, Horner, May, & Vincent, 2008), it may be that fewer elementary schools need such a comprehensive, lengthy program like IY-TCM. One indicator that a training program is needed could be high rates of office discipline referrals or observation of harsh, negative teacher behaviors. IY-TCM principles and strategies could also be taught through individual teacher consultation (Williford & Shelton, 2008), targeting elementary teachers who are struggling to implement PBIS or who have limited effective management strategies and a harsh interaction style. Another potential practice recommendation is to provide IY-TCM for all beginning teachers or as part of pre-service training. Alternately, there are components of IY-TCM which were considered highly valuable by teachers participating in this study (e.g., praise and behavior planning) which could be provided in briefer in-service trainings to all teachers. Such a group learning approach may be particularly beneficial given teachers' documented appreciation for the peer collaboration provided by IY-TCM.

With regard to implications for future research with similar universal teacher professional development programs, matching the program to the setting is one important recommendation. Similar to the practice recommendation noted above, this will likely involve conducting assessments of teachers' skills and/or perceived needs at an individual or school-wide level. Again, given the increasing adoption of PBIS (Spaulding et al., 2008), tools such as the school-wide evaluation tool (Horner et al., 2004) or high rates of office discipline referrals may provide such indication of need. Another potential research implication is that, unlike preschool-aged children for whom more robust benefits from universal teacher training have been observed (Bierman & Torres, 2016), early elementary-aged students who exhibit delayed self-regulation skills (e.g., at-risk students) may need more intensive interventions including direct skills training. Current work by the first author is underway to evaluate the IY Dina Dinosaur small-group Treatment Program as delivered in schools (R305A150169), which will help address this question. Another area where future research is needed is examination of how social-emotional programs may impact executive functioning, which is a foundational cognitive self-regulation skills related to academic competence (Diamond, 2012).

References

- Baker-Henningham, H., Walker, S., Powell, C., & Gardner, J. M. (2009). A pilot study of the Incredible Years Teacher Training Programme and a curriculum unit on social and emotional skills in community pre-schools in Jamaica. *Child: Care, Health and Development*, 35, 624–631. <http://dx.doi.org/10.1111/j.1365-2214.2009.00964.x>.

- 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. <http://dx.doi.org/10.1177/02711214070270030801>.
- Bierman, K. L., & Torres, M. (2016). Promoting the development of executive function through early education and prevention programs. In J. A. Griffin, L. S. Freund, & P. McCordle (Eds.), *Executive function in preschool age children: Integrating measurement, neurodevelopment and translational research* (pp. 299–326). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/147977-014>.
- Briesch, A. M., Hemphill, E. M., Volpe, R. J., & Daniels, B. (2015). An evaluation of observational methods for measuring response to classwide intervention. *School Psychology Quarterly*, 30, 37–49. <http://dx.doi.org/10.1037/spq0000065>.
- Burchinal, J., Zaslow, M., & Tarullo, L. (2016). Quality thresholds, features, and doses in early care and education: Secondary data analyses of child outcomes. *Monographs of the Society for Research in Child Development*, 81, 1–128.
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25, 166–176. <http://dx.doi.org/10.1016/j.jecresq.2009.10.004>.
- Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology*, 43, 61–85. <http://dx.doi.org/10.1016/j.jsp.2004.12.002>.
- Carlson, J. S., Tiet, H. B., Bender, S. L., & Benson, L. (2011). The influence of group training in the Incredible Years Teacher Classroom Management Program on preschool teachers' classroom management strategies. *Journal of Applied School Psychology*, 27, 134–154. <http://dx.doi.org/10.1080/15377903.2011.565277>.
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Conduct Problems Prevention Research Group (1995). *Unpublished technical report, fast track, instruments and measures*. Durham, NC: Duke University. <https://fasttrackproject.org/techrept/t/tsc/>.
- Conduct Problems Prevention Research Group (1999). Initial impact of the Fast-Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*, 67, 631–647. <http://dx.doi.org/10.1037/0022-006X.67.5.631>.
- Conners, C. K. (2001). *Conners' rating scales revised*. North Tonawanda, NY: Multi-Health Systems.
- Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, 98, 665–689. <http://dx.doi.org/10.1037/0022-0663.98.4.665>.
- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Washington, DC: National Staff Development Council.
- Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science*, 21, 335–341. <http://dx.doi.org/10.1177/0963721412453722>.
- Downer, J., Sabol, T. J., & Hamre, B. (2010). Teacher-child interactions in the classroom: Toward a theory of within- and cross-domain links to children's developmental outcomes. *Early Education and Development*, 21, 699–723. <http://dx.doi.org/10.1080/10409289.2010.497453>.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82, 405–432. <http://dx.doi.org/10.1111/j.1467-8624.2010.01564>.
- Eisenberg, N., & Sulik, M. J. (2012). Emotion-related self-regulation in children. *Teaching of Psychology*, 39(1), 77–83. <http://dx.doi.org/10.1177/0098628311430172>.
- Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, 10, 40–67. <http://dx.doi.org/10.1080/19345747.2016.1246632>.
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, 625–638. <http://dx.doi.org/10.1111/1467-8624.00301>.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, 76, 949–967. <http://dx.doi.org/10.1111/j.1467-8624.2005.00889.x>.
- Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes: Conceptualization and measurement. In J. L. Meece, & J. S. Eccles (Eds.). *Handbook of research on schools, schooling, and human development* (pp. 25–41). New York, NY: Routledge.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. Retrieved from <http://www.jstatsoft.org/v42/i08/>.
- Horner, R. H., Todd, A., Lewis-Palmer, T., Irvin, L., Sugai, G., & Bolland, J. (2004). The school-wide evaluation tool (SET): A research instrument for assessing school-wide positive behavior support. *Journal of Positive Behavior Interventions*, 6, 3–12. <http://dx.doi.org/10.1177/10983007040060010201>.
- Hutchings, J., Martin-Forbes, P., Daley, D., & Williams, M. E. (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. <http://dx.doi.org/10.1016/j.jsp.2013.08.001>.
- Ivers, N. M., Halperin, I. J., Barnsley, J., Grimshaw, J. M., Shah, B. R., Tu, K., et al. (2012). Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials*, 13, 120. <http://dx.doi.org/10.1186/1745-6215-13-120>.
- Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, 79, 491–525. <http://dx.doi.org/10.3102/0034654308325693>.
- Jones, S. M., Brown, J. L., Hoglund, W. L., & Aber, J. L. (2010). A school-randomized clinical trial of an integrated social-emotional learning and literacy intervention: Impacts after 1 school year. *Journal of Consulting and Clinical Psychology*, 78, 829–842. <http://dx.doi.org/10.1037/a0021383>.
- Joyce, B. R., & Showers, B. (2002). Student achievement through staff development. In B. Joyce, & B. Showers (Eds.). *Designing training and peer coaching: Our needs for learning*. Association for Supervision and Curriculum Development: Alexandria, VA.
- Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, 48, 361–386. <http://dx.doi.org/10.3102/0002831210382888>.
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, 86, 643–680. <http://dx.doi.org/10.3102/0034654315626799>.
- Lippman, L., Burns, S., & McArthur, E. (1996). *Urban schools: The challenge of location and poverty (report no. NCES 96-184)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York, NY: Wiley.
- Matheson, A. S., & Shriver, M. D. (2005). Training teachers to give effective commands: Effects on student compliance and academic behaviors. *School Psychology Review*, 34, 202–219. Retrieved from <http://libproxy.lib.unc.edu/login?url=http://search.proquest.com/docview/219647773?accountid=14244>.
- McGilloway, S., Hyland, L., Ni Mhaille, G., Lodge, A., O'Neill, D., Kelly, P., ... Donnelly, M. (2010). *Positive classrooms, positive children: A randomised controlled trial to investigate the effectiveness of the Incredible Years Teacher Classroom Management Programme in an Irish context*. Dublin, Ireland: Archways.
- Morris, P., Mattera, S., Castells, N., Bangser, M., Bierman, K., & Raver, C. (2014). *Impact findings from the Head Start CARES demonstration: National evaluation of three approaches to improving preschoolers' social and emotional competence (OPRE report no. 2014-44)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Murray, D., Rosanbalm, K., Christopoulos, C., & Hamoudi, A. (2015). *Self-regulation and toxic stress: Foundations for understanding self-regulation from an applied developmental perspective (OPRE report no. 2015-21)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Murray, D. W., Rosanbalm, K., & Christopoulos, C. (2016). *Self-regulation and toxic stress report 3: A comprehensive review of self-regulation interventions from birth through young adulthood (OPRE report no. 2016-34)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- National Center for Education Statistics (2012). Number and percentage of public school students eligible for free or reduced-price lunch, by state: Selected years, 2000–01 through 2010–11. Retrieved from nces.ed.gov/ipeds/data/ipedsdatatools/tables/d12_046.asp.
- National Center for Education Statistics (2016). Percentage distribution of students enrolled in public elementary and secondary schools, by race/ethnicity: Fall 2003, fall 2013, and fall 2025. Retrieved from https://nces.ed.gov/ipeds/data/ipedsdatatools/tables/d12_046.asp.
- North Carolina Department of Public Instruction (2013a). Facts and figures 2012–13. Retrieved from <http://www.ncpublicschools.org/docs/fbs/resources/data/>

- factsfigures/2012-13figures.pdf.
- North Carolina Department of Public Instruction (2013b). Free and reduced meals application data 2012–13. Retrieved from <http://www.dpi.state.nc.us/fbs/resources/data/>.
- Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011, March). *Teacher classroom management practices: Effects on disruptive or aggressive student behavior. Paper presented at the conference of Society for Research on Educational Effectiveness, Washington, D.C.*
- Pianta, R. C., DeCoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J., ... Howes, C. (2014). Dose–response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Childhood Research Quarterly*, 29, 499–508. <http://dx.doi.org/10.1016/j.jecresq.2014.06.001>.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system*. Baltimore, MD: Brookes Publishing.
- Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal*, 102, 225–238. <http://dx.doi.org/10.1086/499701>.
- Rabiner, D., Coie, J. D., & The Conduct Problems Prevention Research Group (2000). Early attention problems and children's reading achievement: A longitudinal investigation. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 859–867. <http://dx.doi.org/10.1097/00004583-200007000-00014>.
- Rabiner, D. L., Murray, D. W., Schmid, L., & Malone, P. S. (2004). An exploration of the relationship between ethnicity, attention problems, and academic achievement. *School Psychology Review*, 33, 498–509. Retrieved from <http://libproxy.lib.unc.edu/login?url=http://search.proquest.com/docview/219654478?accountid=14244>.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publishing.
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 77, 302–316. <http://dx.doi.org/10.1037/a0015302>.
- Reinke, W. M., Herman, K. C., & Dong, N. (2014, March). *A group randomized evaluation of the Incredible Years Teacher Training Program. Paper presented at the Annual Meeting of Society for Research on Educational Effectiveness, Washington, DC.*
- Reinke, W. M., Stormont, M., Webster-Stratton, C., Newcomer, L. L., & Herman, K. C. (2012). The Incredible Years Teacher Classroom Management Program: Using coaching to support generalization to real-world classroom settings. *Psychology in the Schools*, 49, 416–428. <http://dx.doi.org/10.1002/pits.21608>.
- Roscigno, V. J., & Crowle, M. L. (2001). Rurality, institutional disadvantage, and achievement/attainment. *Rural Sociology*, 66, 268–292. <http://dx.doi.org/10.1111/j.1549-0831.2001.tb00067.x>.
- Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2004). Attentional control and self-regulation. In K. D. Vohs, & R. F. Baumeister (Eds.). *Handbook of self-regulation: Research, theory, and applications* (pp. 284–299). New York: Guilford Press.
- Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher–child relationship during the school transition. *Journal of School Psychology*, 43, 39–60. <http://dx.doi.org/10.1016/j.jsp.2004.11.003>.
- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, 31, 351–380. Retrieved from <http://www.jstor.org/stable/42899983>.
- Spaulding, S., Horner, R., May, S., & Vincent, C. (2008). Evaluation brief: Implementation of school-wide PBIS across the United States. Retrieved from <http://www.pbis.org/blueprint/evaluation-briefs/implementation-across-us>.
- Teachstone (2015). State of CLASS 2015 report. Retrieved from http://info.teachstone.com/hubfs/State_of_CLASS_Final.pdf?t=1473877878567.
- U.S. Public Health Service (2000). *Report of the surgeon general's conference on children's mental health: A national action agenda*. Washington, DC: Department of Health and Human Services (Stock No. 017-024-01659-4; ISBN No. 0-16-050637-9).
- Webster-Stratton, C. (1994). *The Incredible Years Teacher Training series*. Seattle, WA: Incredible Years, Inc.
- Webster-Stratton, C. (1999). *How to promote children's social and emotional competence*. London: Sage.
- Webster-Stratton, C. (2012). *Incredible teachers: Nurturing children's social, emotional, and academic competence*. Seattle, WA: Incredible Years, Inc.
- Webster-Stratton, C., Reid, M. J., & 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. http://dx.doi.org/10.1207/S15374424JCCP3003_2.
- Webster-Stratton, C., Reid, M. J., & 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. http://dx.doi.org/10.1207/S15374424JCCP3301_11.
- Webster-Stratton, C., Reid, M. 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. <http://dx.doi.org/10.1111/j.1469-7610.2007.01861.x>.
- Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. 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. Retrieved from <http://libproxy.lib.unc.edu/login?url=http://search.proquest.com/docview/921366095?accountid=14244>.
- Williford, A. P., & Shelton, T. L. (2008). Using mental health consultation to decrease disruptive behaviors in preschoolers: Adapting an empirically-supported intervention. *Journal of Child Psychology and Psychiatry*, 49, 191–200. <http://dx.doi.org/10.1111/j.1469-7610.2007.01839.x>.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33(2 Suppl), S130–S143. <http://dx.doi.org/10.1016/j.amepre.2007.04.011>.
- Yarrow, A., Ballantyne, R., Hansford, B., Herschell, P., & Millwater, J. (1999). Teaching in rural and remote schools: A literature review. *Teaching and Teacher Education*, 15, 1–13. [http://dx.doi.org/10.1016/S0742-051X\(98\)00036-5](http://dx.doi.org/10.1016/S0742-051X(98)00036-5).