



Evaluation of the Incredible Years Teacher Classroom Management Program in a Regular Norwegian School Setting

Merete Aasheim ¹⁰ a, Charlotte Reedtz ¹⁰ b, Bjørn Helge Handegård Monica Martinussen ¹⁰ and Willy-Tore Mørch Monica Martinussen ¹⁰ b

^aFaculty of Health Sciences, Regional Centre for Child and Youth Mental Health and Child Welfare - (RKBU-North), UiT The Arctic University of Norway, Tromsø, Norway; ^bFaculty of Health Sciences, RKBU-North, UiT The Arctic University of Norway, Tromsø, Norway.

ABSTRACT

This study examined whether the Incredible Years (IY) Teacher Classroom Management (TCM) program implemented as a school-wide preventive intervention at 1st to 3rd grade in a regular school setting reduces the development of problem behavior and improves social competence. Using a quasi-experimental pre-post design, the IYTCM was implemented in 21 schools and compared with 22 matched schools that did not receive the program. A total of 241 1st to 3rd grade teachers and 1518 students aged 6 to 8 years took part in the trial. Mixed-model analyses found small positive effects on changes in students' social competence ($d_w = 0.19$), while effects on change in students problem behavior were less than small ($d_w < 0.20$). When the program is implemented as school-wide universal preventive intervention, results suggest a small preventive impact of the IYTCM program in regular school settings for some of the outcomes measured in the study.

ARTICLE HISTORY

Received 30 June 2017 Accepted 11 April 2018

KEYWORDS

Problem behavior; social competence; school-wide; prevention

Behavioral problems in school are associated with educational and social disadvantages and are one of the most prevalent, severe, and persistent problems that inhibit the realization of students' abilities. Such problems may have both immediate and long-term consequences for the student, such as academic underachievement, mental health problems, school dropout, future unemployment, and general social exclusion (Ford et al., 2012; Scott, Knapp, Henderson, & Maughan, 2001). There are huge costs to the public sector associated with behavioral problems, particularly in the education system (Snell et al., 2013). Measures aimed at preventing and reducing problem behavior and coping issues are central to ensuring that all students experience optimal development and positive learning outcomes at school. Preventive interventions in school settings can be of great importance as public health interventions (Ford et al., 2012). Dysfunctional patterns of family interaction often translate into problems at school, underlining the need to target behavioral problems not only at home but also in day-care and school settings (Drugli & Larsson, 2006; Fossum, Handegård, Martinussen, & Mørch, 2008; Ramsey, Patterson, & Walker, 1990). Students who exhibit problem behavior frequently go off-task, display aggression towards others, or refuse to cooperate, all of which adversely affects their own learning potential as well as that of the students around them (Bartlett, Holditch-Davis, Belyea, Halpern, & Beeber, 2006; Efrati-Virtzer & Margalit, 2009; Moffitt & Scott, 2009). Students' oppositional and negative behavior may be reinforced by teachers' ineffective classroom behavior management practices, where the teacher is trapped into coercive exchanges with the student because of compliance to students' demands (Patterson, Reid, & Dishion, 1992; Webster-Stratton, Reid, & Hammond, 2001; Webster-Stratton, Reid, & Stoolmiller, 2008).

Social competence, emotional self-regulation, and absence of problem behavior are important components of the foundations of interpersonal adjustment and academic success (Drugli, Klökner, & Larsson, 2011; Snyder et al., 2011; Webster-Stratton et al., 2008). In addition to their educational benefits, an essential developmental task for students is to learn how to interact in socially appropriate ways. Social skills are invaluable in almost every interaction that a student encounters in the school environment, and are a prerequisite for academic learning since they involve self-regulation, the ability to give and receive help, and the skills of working with, listening to, and communicating with others. Students who lack these skills are likely to suffer socially, and to develop problem behaviors that impair their academic progress (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Classrooms provide excellent settings for targeting students' behavior, and teachers are natural implementers who can have a significant influence on their behavior (Greenwood, Kratochwill, & Clements, 2008; Reid, Webster-Stratton, & Beauchaine, 2001). However, the risk of developing behavioral problems may be increased in poorly managed classrooms (Conroy, Sutherland, Haydon, Stormont, & Harmon, 2009; Reid et al., 2001; Reinke, Herman, & Dong, 2016; Webster-Stratton et al., 2008).

The Incredible Years (IY) Teacher Classroom Management (TCM) program aims to strengthen teachers' use of evidence-based classroom management strategies in order to reduce early-onset problem behavior and promote students' social competence. The IYTCM program has been the subject of comprehensive empirical examinations in various combinations with the IY parent and IY child programs: for example, in Head Start centers with high-risk students, students from low socioeconomic backgrounds, and in schools that receive a higher level of support in terms of pupil-teacher ratios, special school grants and extra support for students. Previous studies that measured child outcomes have been linked to reductions in conduct problems, aggression, hyperactivity, and antisocial behavior, as well as improvements in on-task behavior, increased prosocial behaviors, and school readiness (Baker-Henningham, Scott, Jones, & Walker, 2012; Baker-Henningham, Walker, Powell, & Gardner, 2009; Reinke et al., 2016; Webster-Stratton et al., 2001; Webster-Stratton et al., 2008; Webster-Stratton, Reid, & Hammond, 2004). A few studies have also evaluated the impact of the IYTCM program as a stand-alone intervention aimed at changing student behavior. McGilloway et al. (2010) found that when teachers increased their use of positive IYTCM classroom management strategies in combination with reduced use of negative classroom management strategies, student behavior and socioemotional adjustment improved, particularly among those considered initially to be at most risk. In addition, Hutchings, Daley, Jones, Martin, and Gwyn (2007) and Hutchings, Martin-Forbes, Daley, and Williams (2013) found significant reductions in the total number of commands (e.g., negative instructions) given to children, which in turn led to an increase in the rate of compliance (e.g., children paid more attention and were more likely to cooperate with their teachers), after the IYTCM intervention.

The Norwegian School Context

Norway has a mandatory school system for children aged 6 to 16. About 633,000 students are enrolled in primary and lower-secondary school (1-10). Of the 8% (68% boys) who receive special education, 39% receive it as part of ordinary classes, and not in segregated settings. About 7% of the students have a first language other than Norwegian and received special education in Norwegian in parallel with their ordinary education. The schools are divided into the categories small (< 200 students), medium (201-350 students) and large (351-780 students) (Statistics Norway, 2017). In primary school, the average ratio of students to teachers is 16:1. Schools are mostly public and free of charge, and the local authorities are responsible for primary and lower-secondary education. The stages are based on a single national curriculum, which is based on the concept of equality, inclusion, and adapted education for all.



The Current Study

Convincing findings have been found for the IYTCM program in 3- to 8-year-old children in various disadvantageous school settings, both in combination with the IY child and/or the IY parent program, and as a stand-alone intervention (Baker-Henningham et al., 2009; Baker-Henningham et al., 2012; Hutchings et al., 2007; Hutchings et al., 2013; McGilloway et al., 2010; Reinke et al., 2016; Webster-Stratton et al., 2001, 2004; Webster-Stratton et al., 2008). This study is one of the first universal preventive evaluations of the IYTCM program implemented as a school-wide intervention in 1st to 3rd grade to students aged 6 to 8 years. The training was delivered simultaneously to the entire first- to third-grade teaching staff and to after-school service staff. We formulated the following hypotheses: the IYTCM program, when provided as a school-wide preventive intervention in a regular school setting would (1) reduce the development of problem behavior and (2) improve students' social competence. Group differences in the level of change in problem behavior and social competence that favored the students in the IYTCM group were anticipated.

Methods

Participants

In connection to this study, the municipalities that had previously implemented the IY parenting program were invited by IY Norway to participate in the study and to implement the IYTCM program. Employees in the education agencies were trained as IYTCM program group leaders, and informed the schools about the implementation and research study of the IYTCM program. Recruitment continued through five consecutive years, from fall 2009 to fall 2013. In all, 24 municipalities implemented the IYTCM program; 25 schools from these municipalities applied to IY Norway to implement the program and participate in the study. As a part of the study inclusion criteria, readiness for program implementation with approval from at least 80% of the school staff, as well as agreement with school-wide implementation in the 1st to 3rd grades, needed to be met and 21 of the 25 schools that applied satisfied the inclusion criteria. The IYTCM training was provided free of charge. Four schools that did not met the predefined inclusion criteria and were allocated to the comparison group, were offered IYTCM implementation immediately the year after participation. In order to minimize contamination of the program, 19 schools were recruited to the comparison group from municipalities that had not implemented IY. These schools were offered a modest financial compensation for not receiving the IYTCM training immediately. On request, the municipalities and schools were given implementation support from IY Norway after the study period ended. The comparison group was matched with the IYTCM group according to geographic location and school size. For the 43 schools included at pre-assessment, the mean class size was 19.7 (SD = 8.8). None of the 43 schools were actively attending or had attended any other evidence-based school behavior intervention programs during the previous year. The flow of participants through each stage of the study is illustrated in Figure 1.

The number of teachers in the 1st to 3rd grades was 567. One teacher per class who was in daily contact with the students was asked to participate as respondent. These resulted in 241 teachers; 139 in the intervention and 102 in the comparison group. The number of students in the 1st to 3rd grades was 3331. In order to reduce data dependency and to maximize the effective sample size, as well as to limit teacher burden, a statistician, who was blind to the characteristics of the schools (3rd author, BHH), randomly selected seven students per class for the assessment. For example, if a class consisted of 21 students, a random number sequence list from 1-21 was generated electronically. Thereafter each teacher matched the first seven random numbers from the list with the student's alphabetical order. This resulted in 829 randomly selected students in the intervention and 689 randomly selected students in the comparison group. Teachers who participated as respondents received a small financial compensation for the time they spent on completing the questionnaires. A sub-sample of 83 students (6%) scored equal to or above the



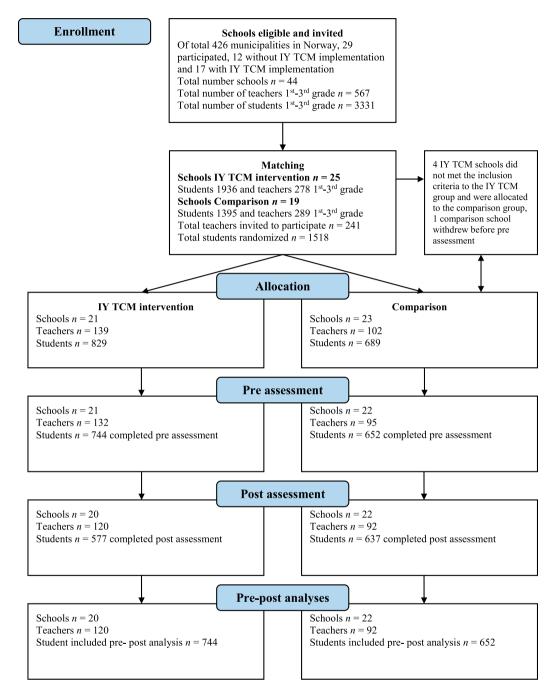


Figure 1. Flow diagram.

90th percentile on the Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R) scale (> 144), which is equal to the clinical range. The findings for the high-risk students are presented in Kirkhaug et al. (2016). Table 1 presents demographic information for the schools, teachers, and students in the study. Apart from one significant difference between groups in terms of student ethnicity (p < .001), none of the other demographical variables showed significant group differences at the .05 level.

Table 1. Demographic information for schools, teachers and students at baseline.

	IYTCM	Comparison	Total	
Schools (n)	21	22	43	
School size large (351–780 students)	4	2	6	
School size medium (201–350 students)	6	4	10	
School size small (< 200 students)	11	16	27	
Class size M (SD)	20.82 (6.85)	18.48 (10.55)	19.74 (8.84)	
Teacher responders (n)	132	95	227	
Teacher's age in years M (SD)	40.94 (11.86)	45.19 (10.31)	42.75 (11.26)	
Work experience in year M (SD)	11.73 (9.11)	15.48 (8.68)	13.37 (9.05)	
Educated as teacher n (%)	123 (93.2)	87 (91.6)	210 (92.5)	
Female teacher n (%)	115 (87.1)	84 (88.4)	199 (87.7)	
Students (n)	744	652	1396	
Girls <i>n</i> (%)	355 (47.7)	297 (45.6)	652 (46.7)	
Age M (SD)	7.22 (0.86)	7.30 (0.87)	7.26 (0.87)	
Non-Norwegian n (%)*	64 (8.6)	13 (2.0)	77 (5.5)	
Special education n (%)	67 (9.0)	72 (11.0)	139 (10.0)	
High-risk student's ^a <i>n</i> (%)	45 (6.1)	38 (5.8)	83 (6.0)	

Note: IYTCM = Incredible Years Teacher Classroom Management.

Procedure

This study had a quasi-experimental pre-post design with a continuous enrollment of intervention and comparison schools. Prior to the pre-assessment and the first IYTCM training, information about the IYTCM program and data collection procedures was presented to teachers and staff. Pre-assessment (Time 1) took place during the fall, one-to-three weeks ahead of the first IYTCM training, and post-assessment (Time 2) was carried out in the spring of the same academic year, one-to-three weeks after the final IYTCM training. The period between the two assessments was typically between eight and nine months. Parents were informed about the IYTCM program and the study, including the data collection procedures, through written information or verbally during parent meetings, and were requested to consent to their children's participation. Provided there was parental consent, the teacher filled out questionnaires about the student. The questionnaires were only available in Norwegian. Students whose parents did not speak Norwegian were excluded. The study population included 5.5% non-Norwegians. In order to ensure confidentiality, the names of the schools, teachers, and students were anonymized using ID-codes. Parents could withdraw their child from the study at any time without further explanation. The questionnaires were returned in pre-paid envelopes or completed using the Internet survey tool Quest Back.

The study was approved by the Regional Committee for Medical and Health Research Ethics, Norway (Approval/reference number: 200803705-7/MGA006/400).

The Intervention

The IYTCM program was developed as a preventive intervention designed to strengthen teacher's classroom management strategies. This in order to reduce early-onset problem behavior, aggression, and non-cooperation in students and to promote students social competence and school-readiness. A basic premise of IYTCM training is that positive teacher-student interaction precedes effective teaching strategies, and that teachers' attention should be directed far more frequently to positive student behaviors in classroom environments than to negative behaviors (Webster-Stratton, 2012). Six topics are covered, with one workshop for each topic, in which each workshop builds upon the content of the previous one, and are delivered as follows: (1) building positive relationships between students and teacher; (2) teacher attention, coaching, encouragement, and praise; (3) motivating students through incentives; (4) reducing inappropriate behavior – ignoring and redirecting; (5) reducing inappropriate behavior – follow-through with consequences; and (6) emotional

^aScore of 144 or higher on SESBI-R Intensity.

^{*}p < .05.

regulation, social skills, and problem solving. Two experienced and qualified group leaders trained the teachers and staff simultaneously in groups (20 in each group), through six full-day workshops over an eight-tonine- month period (about one workshop per month), 42 hours in total. The training started in the fall and was completed during the following spring. Teachers were instructed to practice the principles of the program during the month following each training session and to report on their experiences at the start of the following session. The group leaders provided teachers and staff with guidance during the month after each workshop. As part of the training, the textbook How to Promote Social and Emotional Competence in Young Children (Webster-Stratton, 1999) was provided to teachers and staff. Fidelity in training was promoted by means of checklists completed by both group leader- and teacher, as instructed in the program manual, in order to ensure evidence-based implementation of the program (Webster-Stratton, 2011). Teachers also completed a user satisfaction questionnaire at the end of the training.

To become a qualified group leader, a 21-hour mandatory IYTCM training course provided by IY Norway had to be completed. A higher education qualification (bachelor's or master's degree) in teaching, special education, psychology, health, or social studies, in addition to suitable personal characteristics, were also required. To maintain approval as a qualified group leader, the group leaders had to deliver the training program at least once or twice (or in one or two schools, depending on school size) per year on average, which also was the requirement before they could complete the training for this study. All the group leaders were trained by the same two IYTCM mentors (certified in both the Parenting and the TCM program by the program originator) and supervised by the same two mentors through the data-acquisition period.

Measures

The SESBI-R (Eyberg & Pincus, 1999) was used to evaluate the current frequency and severity of various student behaviors. The 38-item scale describes common behavior problems rated by teachers, such as "teases or provokes other students"; "has difficulty staying on task"; and "fails to listen to instructions." On the Intensity Scale, the frequency of behaviors is rated using a seven-point Likert scale: 1 = never, 2-3 = seldom, 4 = sometimes, 5-6 = often, and 7 = always. On the Problem scale, teachers assess whether or not the behavior is currently a problem for the teacher using a yes-no (1-0) scale. The scores were summed across all items on both the Intensity scale (ranging from 38 to 266) and the Problem scale (ranging from 0 to 38). Clinical cut-off values were as provided by Kirkhaug, Drugli, Mørch, and Handegård (2012). Cronbach's alphas for the baseline data were .97 for the Intensity scale and .95 for the Problem scale.

The Teacher Report Form (TRF: Achenbach & Rescorla, 2001) measures different behavioral difficulties, including the Aggression, Attention problem and Academic Performance subscales employed in this study. The TRF Academic Performance scale evaluates students' overall and current academic functioning, where teachers assess the student in six different academic subjects of the teacher's choosing, rating them on a scale from 1-5 (1 = far below average to 5 = far above average). The average of these scores constitutes the TRF Academic Performance score. In addition, teachers were asked to rate the degree of emotional and behavioral problems observed in students, either currently or during the past two months, using a 0-2 scale (0 = not true as far as you know; 1 = somewhat or sometimes true; 2 = very true or often true). For the TRF scores, Cronbach's alpha was calculated on the baseline data in this study. The alphas for the TRF Internalizing subscales in this study were 0.79 (Anxious/Depressed), 0.72 (Withdrawn/Depressed), and 0.53 (Somatic Complaints), while for the TRF Externalizing subscales, the alphas were 0.94 (Aggressive Behavior) and 0.83 (Rule-Breaking behavior). For the TRF Attention Problem, the alpha was 0.91, for Social Problems 0.73, and for Thought Problems 0.77. Mean test-retest reliability was 0.90 across all TRF scales for US samples by Achenbach and Rescorla (2001).

We used the Social Skills Rating System (SSRS: Gresham & Elliott, 1990) version for the elementary school teacher, which contains 30 items, including subscales for Cooperation, Assertion and SelfControl. Cooperation comprises behaviors such as helping others, sharing, and complying with rules, whereas assertion includes initiating behaviors, such as asking others for information and responding to the actions of others. Behaviors that emerge in conflict situations, such as responding appropriately to teasing, and behaviors that arise during non-conflict situations, such as taking turns and compromising, are including in the Self-Control subscale. The teachers rated how often each social skill occurred using a 1-4 scale (1 = never to 4 = very often). The alphas were calculated using baseline data and were found to be .91 for the Cooperation subscale, .88 for the Assertion subscale, and .87 for the Self-Control subscale. The SSRS total score (ranging from 0-90) was computed across all items and used in the analysis. The SSRS is a well-validated assessment tool and the test-retest reliability of the SSRS has been found to be high (Elliott, Gresham, Freeman, & McCloskey, 1988; Ogden, 2003).

Statistics

Before the main analysis was conducted, independent t-tests and Pearson's chi-squared tests were used to test for group differences on demographic variables. The data were hierarchically organized, with students (Level 1) nested within teachers (Level 2). Linear mixed model (LMM) analysis was used to test for group differences on baseline scores, and for group differences in change in student behavior from pre- to post-assessment, as this is a suitable method for analyzing hierarchical data. Intra-class correlations (ICCs) were calculated to estimate the degree of dependency within-teacher that this clustering causes. Intra-class correlation calculations were based on change scores, since change scores were used as dependent variables in the main analyses.

In order to deal with missing data, multiple imputation was used for the analyses, creating 20 complete sets of data. The imputation was performed on both pre- and post-assessment student variables. The imputation model included demographic variables and all relevant student variables. In the imputation of missing pre- and post-data, all other pre- and post-student variables were used as predictors. Under the assumption of data missing at random, performing multiple imputation of data is an appropriate and flexible way of handling missing data and was therefore done in order to ensure that the pre- and post-analyses reflected the entire student population that participated in this study (Stuart, Azur, Frangakis, & Leaf, 2009). Effect sizes (d_w) were computed as standardized group differences in pre-post mean change using the pooled within-cluster sample standard deviation (Hedges, 2007). A significance level of .05 was adopted for all tests.

Results

Attrition

At pre-assessment, 227 (94%) of 241 teachers participated as respondents and 1396 (92%) of 1518 possible students were included. Drop-out was due to lack of parental consent or delayed arrival of consent forms from parents, as well as insufficiently completed questionnaires, and amounted 7 teachers and 85 students in the intervention, and 7 and 37, respectively, in the comparison group. In both pre- and post-assessments, 212 (88%) teachers and 1214 (80%) students were included. A different dropout pattern at post-assessment was found between the conditions, in that, 167 students in the intervention and 15 students in the comparison group had missing data. Drop-out in the intervention group was due to withdrawal of one school, which included 7 teachers and 49 students, and teachers on leave of absence or changing their jobs, these last included 5 teachers and 28 students. Drop-outs were also due to missing replies, incomplete questionnaires, or protocol errors, which resulted in a further 90 students missing in the intervention group, and 3 teachers and 15 students in the comparison group (see Figure 1). When students who had missing data at post-assessment were compared with students who had both pre- and post-assessment data, students who had missing data at post-assessment differed significantly on SESBI-R Intensity (t = -3.36, p = .02), SESBI-R Problem (t = -2.24, p = .03), and TRF Attention (t = -3.02, p = .003) at pre-assessment. However, no



interaction effects between the intervention group and the dropout group on outcome variables at pre-assessment were found, indicating that pre-assessment differences in the dropout groups were similar in the intervention group and the control group.

Group effects in students' problem behavior measured with SESBI-R and TRF.

There were no significant differences between the conditions at pre-assessment on SESBI-R scores. For group effects in student problem behavior measured with SESBI-R, significant group differences in pre-post change on SESBI-R Intensity and on SESBI-R Problem were found, although the effect sizes were small. Calculations of the ICCs suggested that 22% of the variance on SESBI-R Intensity and 14% of the variance on SESBI Problem might be due to clustering effects among teachers (see Table 2). When we tested for moderating effects of the level of behavior problems (high/low), a significant interaction between treatment group and high-risk status on SESBI-R Intensity was detected. Study of this interaction showed a significantly higher treatment effect for high-risk students compared to those not in the high-risk group (9.9 point pre-post change difference, t = -2.13, p = .03).

For TRF scores, there were no significant differences between the conditions at pre-assessment. For group effects in students' behavioral difficulties measured with TRF, there was a significant group difference in pre-post change on TRF Total. Examination of the TRF subscales revealed significant group differences in pre-post change on Attention Problems, but not in change on Aggressive Behavior or on Academic Performance. For the TRF scores, the effect sizes were small. The ICC calculations on the TRF change scores ranged from 0.6 to 0.14 (see Table 2). The change scores on the TRF Attention Problem also correlated highly with SESBI-R Intensity scores (r = .65).

Group effects on students' social competence measured with SSRS.

There were significant differences between the conditions on the subscales SSRS Cooperation and Self-Control at pre-assessment (see Table 2). For group effects in social competence, a significant

Table 2. Descriptive statistics at pre- and post-test, and results of multilevel analyses examining group differences in pre-post change scores and effect sizes (d_w) .

	TCM intervention ^a		Comparisona		Baseline ^b		Intervention effects ^b	
	Pre (n = 557–722)	Post $(n = 442-577)$	Pre (n = 548–627)	Post (n = 551–634)	Pre		Pre-post	
	M (SD)	M (SD)	M (SD)	M (SD)	t	ICC	t	d_w
SESBI-R								
Intensity in behavior	79.85 (34.89)	78.83 (33.56)	77.45 (33.20)	79.47 (34.02)	-0.89	0.22	-3.021**	0.08
Behavior is a problem	3.29 (6.48)	3.11 (6.39)	2.75 (6.32)	2.90 (6.60)	-0.96	0.14	-2.047*	0.09
TRF								
Aggressive behavior	2.10 (4.91)	2.16 (4.73)	1.70 (4.26)	1.80 (4.43)	-1.13	0.06	-1.678	0.08
Attention problems	5.30 (7.78)	5.15 (7.72)	4.72 (7.37)	5.04 (7.92)	-1.27	0.11	-2.651**	0.08
Academic performance	3.16 (0.50)	3.22 (0.55)	3.19 (0.53)	3.18 (0.59)	0.67	0.09	1.882	0.08
Total problems	10.54 (16.58)	11.50 (16.41)	9.44 (15.47)	9.97 (16.32)	-1.40	0.14	-2.137*	0.09
SSRS								
Cooperation	29.54 (5.90)	30.46 (5.94)	30.45 (7.04)	30.30 (6.27)	2.01*	0.19	2.941**	0.17
Assertion	25.95 (4.95)	27.22 (4.85)	26.40 (4.85)	27.23 (4.73)	0.74	0.40	0.960	0.11
Self-control	27.29 (4.80)	28.39 (5.28)	28.09 (5.10)	28.43 (5.28)	1.98*	0.34	2.389*	0.20
Social skills total	82.80 (13.03)	86.00 (13.13)	84.90 (13.80)	85.95 (14.00)	1.75	0.30	2.403*	0.19

Note: TCM = Incredible Years Teacher Classroom Management, SESBI-R = Sutter-Eyberg Student Behavior Inventory-Revised, TRF = Teacher Report Form, SSRS = Social Skills Rating System, ICC = Intra-class correlations. dw = Effect sizes were computed using the pooled within-treatment groups' standard deviation of the cluster means (pre assessments scores).

All covariates gender, grade, ethnicity, special education, how well the teacher knew the student, number of hours the teachers taught the student each week, and number of students in each class were statistically accounted for in the different multilevel analyses.

^aoriginal data.

bimputed data.

^{*}p < .05.

^{**}p < .01.

group difference in pre-post change on SSRS Total was found. The SSRS subscale results showed significant group differences in change on SSRS Cooperation and SSRS Self-control. The preventive effect sizes on the SSRS scores were higher than the SESBI-R and the TRF scores, although, the effects were in the small range ($d_w = 0.11-0.20$). Calculations of the ICCs on SSRS change scores varied from 0.19 to 0.40 (see Table 2). Testing for moderating effects of the level of behavior problems, as well as of grade and gender, revealed a significant interaction between treatment group and grade on the SSRS Total (F = 3.26, p = .04). While this interaction revealed a significantly larger treatment effect in 2nd grade compared to 3rd grade (t = -2.55, p = .01), the treatment effects in 1st grade compared to the 2nd and 3rd grades were not significant. For further details about group differences in prepost changes, and the sizes of the effect, see Table 2.

Discussion

The aim of this quasi-experimental pre-post control group study was to evaluate the universal preventive impact of the IYTCM program. The program was implemented as a universal school-wide preventive intervention in 21 schools, with the aim of reducing the development of problem behavior and improving social competence in students. Our first hypothesis, that the IYTCM program would reduce the development of problem behavior, was to some extent supported. There was a significant group difference in favor of the IYTCM group on the intensity of problem behavior, and a moderation analysis indicated that the program had a larger effect on students with elevated intensity scores at pre-assessment than on students with lower scores at pre-test. The total score for change in social, emotional, and behavioral problems, including the scores for the subscale TRF Attention Problem, was significant, in favor of the IY TCM group. The SESBI-R is a general measure of behavior problems, although Kirkhaug et al. (2012) suggested that the SESBI-R has two separate, measurable factors, of which the first reflects oppositional behavior and the second attentional difficulties. Hence, the finding for TRF Total corresponds with the findings for the SESBI-R Intensity scores.

The second hypothesis, that training teachers in the IY TCM program would improve student social competence, was also partially supported. However, the pre-assessment scores on SSRS Cooperation and Self-Control were less favorable in the IYTCM group than the comparison group, while the scores for both groups were almost equal at post-assessment. The significance of the effects of the IYTCM on SSRS Cooperation and SSRS Self-Control might be questioned. The fact that the IYTCM schools applied for implementation in the program and participation in the study may explain this finding. The self-recruitment may be due to a need to address existing but general issues, and this may have led to a higher level of awareness when their students' behavior at pre-assessment was being evaluated. Whether the difference in change for SSRS Total is due to an actual effect of the program on social skills for students in the IYTCM group or to a regression towards the mean, may therefore be a matter of interpretation (Barnett, Van Der Pols, & Dobson, 2015; Shadish, Cook, & Campbell, 2002).

However, when the intensity of problem behavior is reduced among the high-risk group of students in a class, is seems fair to suggest that the whole class may have profited from a reduction in problem behavior in some students. The treatment effect on the intensity of problem behavior in the high-risk group of students was about 10 points higher than among the rest of the students. These changes may have positively affected the high-risk group of students own potential to learn as well as that of the students around them. Collectively, the students who dropped out between pre- and postassessment scored less favorably on SESBI-R intensity, SESBI Problem, and TRF Attention at preassessment. This could have reduced the overall effects of the intervention, since changes in the study were larger among students with elevated intensity scores at pre-assessment compared to those with lower scores at pre-test. Additionally, the main outcomes were more evident when the IYTCM program was evaluated in a younger kindergarten cohort within the same study as ours where the children's mean age was 4.4 (SD = 0.9) (Fossum, Handegård, & Drugli, 2017), as compare to the mean age for students in our study which was 7.3 (SD = 0.9). In the Fossum et al. (2017) study,

significant preventive effects on change in children's problem behavior, aggressive behavior, internalization, and attention problems, as well as improvement in social competence, were found.

Furthermore, unlike the IY Dina Dinosaur Social Skills and Problem Solving Curriculum, the IYTCM training covers the socioemotional curriculum in only one of the training days. The effects of the TCM program have been shown to be more explicit when the TCM training is carried out in combination with the Dina Dinosaur training, both in general, and especially for children who initially scored high on problem behavior and low on social competence (Baker-Henningham et al., 2009; Webster-Stratton et al., 2008). However, when interventions are examined in disadvantageous school settings, the effects are often greater than effectiveness trials that are carried out in normal school settings (Durlak et al., 2011; Weare & Nind, 2011). Most Norwegian students behave well (Nordahl, Mausethagen, & Kostøl, 2009), and students in this study scored within the typical range of Norwegian children on problem behavior. Mean scores in the Kirkhaug et al. (2012) study on SESBI-R Intensity (M = 83.8, SD = 38.6), and mean scores in Larsson and Drugli (2011) study on TRF Total (M = 15.5, SD = 19.0) were less favorable, than compared to the mean SESBI-R Intensity scores (M = 78.7, SD = 34.1) and the mean TRF Total scores (M = 10.0, SD = 16.0) in our study. The implementation of the program was also naturalistic and with restricted control regarding fidelity of the intervention. Large effect sizes were therefore not to be expected. Our effect sizes were in the range of 0.08 to 0.20, yet even small effects in statistical terms may lead to improvements in the ability of students to engage in positive relationships with their schools, which in itself is known to be a protective factor against long-term behavioral problems. Thus, these small effects may have practical importance for many students in the long run (Weare & Nind, 2011). Higher levels of fidelity and implementation monitoring may improve the findings, which could in turn strengthen the preventive effects of IYTCM in regular school settings in Norway (Durlak & DuPre, 2008; Greenberg & Abenavoli, 2017; Sørlie & Ogden, 2015).

Limitations

This study has a number of limitations that need to be taken into consideration. First, the implementation of the IYTCM program was dependent on locally available and qualified group leaders. Since extensive predefined criteria for the implementation of the TCM program, as recommended by IY Norway, had to be fulfilled before study participation, schools needed to apply to IY Norway for program implementation. The recruitment of intervention schools was therefore based entirely on applications from individual schools. This meant that a truly randomized, controlled trial was difficult to achieve. In order to minimize threats to validity such as diffusion (contamination), recruitment of schools to the comparison group was carried out in municipalities that lacked IY implementation. The schools in the intervention group received the IYTCM program free, whereas the schools in the comparison group received a minor financial compensation instead of implementation of the program. The situation for the comparison schools was therefore different from that of the IYTCM schools. Slightly elevated pre-scores in the intervention group suggest that some of the schools that sent a request for implementation of the program (self-recruitment) may have realized that they had issues with student behavior and that they could benefit from implementing the IYTCM program. Hence, a potential selection threat due to the sampling strategy may have affected our results. An alternative design might have been a step-wedge design, but that was discarded because it would have resulted in an excessive burden for the participants.

Second, the implementation process was partly in the hands of the local authorities involved, and access to information about the fidelity was not accessible due to practical limitations. The Norwegian Directorate of Health funds the IY Norway, and the authorities meet expenses in connection with organizing curriculum, groups, and training of group leaders. At the time of the study, the fundraiser wished clear boundaries to exist between the implementation of IYTCM in Norway and its research projects, in order to facilitate the independence of research and implementation. This made



it difficult to collect valid data from the implementation process, and we cannot know for certain whether the program was delivered in a less than optimal manner than required by the manual.

Third, it is a significant limitation that teachers were the only source of reports of student behavior. Changes may therefore reflect teachers' perceptions rather than actual changes in behavior. However, teacher observations of students in the classroom or in the school environment may reveal difficulties that cannot be observed otherwise or elsewhere. Teachers are also able to compare their students with other students of the same age and developmental level, and are important informants regarding how well students function at school (Lurie, 2006). Nevertheless, observational data would have improved the robustness of the study and the findings.

Fourth, a different dropout pattern between the intervention and comparison conditions was found. However, interaction effects between the intervention group and the dropout group on outcome variables at pre-assessment were not confirmed. The situation of the comparison schools was different from that of the IYTCM schools, which may have encouraged more sustained participation at post-assessment. In order to compensate for the missing data, multiple imputation was used to ensure that the pre-post analyses reflected the whole of the student population that participated in this study. Analyses performed on imputed data are relatively stringent, and our tests confirmed the results of the LMM analyses of the original data, which improves the generalizability of the findings (Stuart et al., 2009).

Finally, there was no long-term follow-up in this study and, therefore, no evidence as to whether changes in student behavior would be sustained in the future. Furthermore, the short pre-post intervention period may have limited the opportunity for teachers to implement everything they had learned from the IYTCM training. Previous research suggests that behavioral changes realized through classroom interventions may take longer to develop than those achieved in clinical settings; thus, preventive school-wide interventions may need to be implemented consistently over time in order to produce more convincing outcomes (Sørlie & Ogden, 2015; Weare & Nind, 2011).

Conclusion and Implications for School Practice

The findings from this study may provide important implications for promoting effective classroom environments in school. The IYTCM program was delivered as a school-wide universal preventive intervention simultaneously to the entire group of students with varying degrees of risk and within a limited period of time, hence, large effect sizes were not expected (Greenberg & Abenavoli, 2017). Differential effects from universal preventive school interventions may also be due to differences in implementation quality (Sørlie & Ogden, 2015), which may explain the lack of more robust findings in our study. To ensure positive program effects and sufficient implementation support, continuous monitoring of factors that contribute to sustained implementation quality, as well as strategies to develop effective partnerships between educational practitioners and local authorities (e.g., the local Educational and Psychological Counseling Service), are needed. Decisions by politicians and school administrators on issues regarding the implementation of evidence-based universal preventive interventions in schools are therefore an important issue. Findings in the present study may suggest that the IYTCM program delivered as a school-wide universal preventive intervention provides an opportunity to influence all students effectively, including students initially most at risk for developing problem behavior, compared to interventions that address only a limited group of teachers, classes, or students (Durlak et al., 2011; Greenberg & Abenavoli, 2017).

Future Research

The quality of teacher-student relationships has been shown to have a significant influence on students' learning and to play an important role in their functioning, both academically and socially (Baker, 2006; Drugli et al., 2011). Moreover, teachers' involvement with parents and parents' ability to collaborate with teachers have also been shown to be important predictors of student functioning

and achievement at school (Desforges & Abouchaar, 2003; Webster-Stratton et al., 2008; Wyrick & Rudasill, 2009). The IYTCM training directly targets teachers rather than students; thus, our findings may reflect the reasonable assumption that changes in student behavior result from changes in teacher behavior. Whether training teachers in the IYTCM program led to change in teacher-student relationships and parent involvement as well as to changes in teacher behavior after implementation of the IYTCM program will be examined in future analyses. Prevention effects often emerge after some delay; clearly, long-term follow-up is required in future research in order to detect any sustained preventive impact of the program as well as whether enhanced implementation quality would improve effects of the program.

Acknowledgements

Thanks goes to Anne-Kari Johnsen, Marte Gamst, and Oddbjørn Løndal for their contribution to the data acquisition. In addition, the authors are grateful to the students and families, school principals, teachers, and all other school personnel from the 43 participating schools who gave their time to the study.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

The study was funded by UiT The Arctic University of Norway and the Norwegian Directorate of Health.

ORCID

Merete Aasheim http://orcid.org/0000-0002-6269-2677 Charlotte Reedtz http://orcid.org/0000-0001-8480-2737 Monica Martinussen b http://orcid.org/0000-0002-0937-1479

References

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for ASEBA school-age forms and profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Baker, J. A. (2006). Contributions of teacher-child relationships to positive school adjustment during elementary school. Journal of School Psychology, 44, 211-229. doi.org/10.1016/j.jsp.2006.02.002

Baker-Henningham, H., Scott, S., Jones, K., & Walker, S. (2012). Reducing child conduct problems and promoting social skills in a middle-income country: Cluster randomized controlled trial. British Journal of Psychiatry, 201, 101-108. doi.org/10.1192/bjp.bp.111.096834

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. doi.org/10.1111/j.1365-2214.2009.00964.x

Barnett, A. G., Van Der Pols, J. C., & Dobson, A. J. (2015). Correction to: Regression to the mean: What it is and how to deal with it. International Journal of Epidemiology, 44, 1748-1748. doi.org/10.1093/ije/dyv161

Bartlett, R., Holditch-Davis, D., Belyea, M., Halpern, C. T., & Beeber, L. (2006). Risk and protection in the development of problem behaviors in adolescents. Research in Nursing & Health, 29, 607-621. doi.org/10.1002/nur.20163

Conroy, M., Sutherland, K., Haydon, T., Stormont, M., & Harmon, J. (2009). Preventing and ameliorating young children's chronic problem behaviors: An ecological classroom-based approach. Psychology in the Schools, 46, 3–17. doi. org/10.1002/pits.20350

Desforges, C., & Abouchaar, A. (2003). The impact of parental involvement, parental support and family education on pupil achievements and adjustment: A literature review. Nottingham, UK: DfES publications. Retrieved from https:// www.nationalnumeracy.org.uk/sites/default/files/the_impact_of_parental_involvement.pdf

Drugli, M. B., Klökner, C., & Larsson, B. (2011). Do demographic factors, school functioning, and quality of studentteacher relationships as rated by teachers predict internalising and externalising problems among Norwegian school children? Evaluation & Research in Education, 24, 243-254. doi.org/10.1080/09500790.2011.626033



- Drugli, M. B., & Larsson, B. (2006). Children aged 4-8 years treated with parent training and child therapy because of conduct problems: Generalisation effects to day-care and school settings. *European Child & Adolescent Psychiatry*, 15, 392–399. doi.org/10.1007/s00787-006-0546-3
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350. doi.org/10.1007/s10464-008-9165-0
- 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. doi.org/10.1111/j.1467-8624.2010.01564.x
- Efrati-Virtzer, M., & Margalit, M. (2009). Students' behaviour difficulties, sense of coherence and adjustment at school: Risk and protective factors. European Journal of Special Needs Education, 24, 59–73. doi.org/10.1080/08856250802596758
- Elliott, S. N., Gresham, F. M., Freeman, T., & McCloskey, G. (1988). Teacher and observer ratings of children's social skills: Validation of the social skills rating scales. *Journal of Psychoeducational Assessment*, 6, 152–161. doi.org/10. 1177/073428298800600206
- Eyberg, S. M., & Pincus, D. (1999). Eyberg child behavior inventory (ECBI) and Sutter-Eyberg student behavior inventory-revised (SESBI-R). Lutz, FL: Psychological Assessment resources.
- Ford, T., Edwards, V., Sharkey, S., Ukoumunne, O. C., Byford, S., Norwich, B., & Logan, S. (2012). Supporting teachers and children in schools: The effectiveness and cost-effectiveness of the Incredible Years teacher classroom management programme in primary school children: A cluster randomised controlled trial, with parallel economic and process evaluations. *BMC Public Health*, 12, 719–719. doi.org/10.1186/1471-2458-12-719
- Fossum, S., Handegård, B. H., & Drugli, M. B. (2017). The Incredible Years teacher classroom management programme in kindergartens: Effects of a universal preventive effort. *Journal of Child and Family Studies*, 26, 2215–2223. doi.org/10.1007/s10826-017-0727-3
- Fossum, S., Handegård, B. H., Martinussen, M., & Mørch, W.-T. (2008). Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents. *European Child & Adolescent Psychiatry*, 17, 438–451. doi.org/10. 1007/s00787-008-0686-8
- 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. doi.org/10.1080/19345747.2016.1246632
- Greenwood, C. R., Kratochwill, T. R., & Clements, M. (2008). School-wide prevention models: Lessons learned in elementary schools. New York, NY: The Guilford Press.
- Gresham, F. M., & Elliott, S. N. (1990). Social skills rating system manual. Circle Pines, MN: American Guidance Service.
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32, 341–370. doi.org/10.3102/1076998606298043
- Hutchings, J., Daley, D., Jones, K., Martin, P., & Gwyn, R. (2007). Early results from developing and researching the Webster-Stratton Incredible Years teacher classroom management training programme in North West Wales. *Journal of Children's Services*, 2, 15–26. doi.org/10.1108/17466660200700023
- 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. doi.org/10.1016/j.jsp.2013.08.001
- Kirkhaug, B., Drugli, M. B., Handegård, B. H., Lydersen, S., Aasheim, M., & Fossum, S. (2016). Does the Incredible Years teacher classroom management training programme have positive effects for young children exhibiting severe externalizing problems in school? A quasi-experimental pre-post study. *BMC Psychiatry*, 16, 1–11. doi.org/10.1186/s12888-016-1077-1
- Kirkhaug, B., Drugli, M. B., Mørch, W.-T., & Handegård, B. H. (2012). Teacher report of children's problem behavior on the Sutter-Eyberg student behavior inventory-revised (SESBI-R) in a Norwegian sample of preschool and school children. *Scandinavian Journal of Educational Research*, 56, 139–153. doi.org/10.1080/00313831.2011.568672
- Larsson, B. O., & Drugli, M. B. (2011). School competence and emotional/behavioral problems among Norwegian school children as rated by teachers on the teacher report form. Scandinavian Journal of Psychology, 52, 553–559. doi.org/10.1111/j.1467-9450.2011.00889.x
- Lurie, J. (2006). Teachers' perceptions of emotional and behavioral problems in 6-12-year-old Norwegian school children. Trondheim: Barnevernets utviklingssenter i Midt-Norge.
- 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 (short-term outcomes). Dublin, Ireland: Archways. Retrieved from http://www.archways.ie/fileadmin/user_upload/Files/Documents/Incredible_Years_Ireland_Study_-_Combined_12_Month_Follow_On_Report.pdf



- Moffitt, T. E., & Scott, S. (2009). Conduct disorders of childhood and adolescence. In M. Rutter, D. V. Bishop, D. S. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), Rutter's child and adolescent psychiatry (pp. 543-564). London: Blackwell. doi.org/10.1002/9781444300895.ch35
- Nordahl, T., Mausethagen, S., & Kostøl, A. K. (2009). Skoler med liten og stor forekomst av atferdsproblemer. En kvantitativ og kvalitativ analyse av forskjeller og likheter mellom skolene [Schools with low and high prevalence of behavior problems: A qualitative and quantitative analysis of school differences]. (Rapport nr. 3 - 2009). Hamar: Høgskolen i Hedmark.
- Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. Scandinavian Journal of Educational Research, 47, 63-76. doi.org/10.1080/00313830308605
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). Antisocial boys. Eugene, OR: Castalia.
- Ramsey, E., Patterson, G. R., & Walker, H. M. (1990). Generalization of the antisocial trait from home to school settings. Journal of Applied Developmental Psychology, 11, 209-223. doi.org/10.1016/0193-3973(90)90006-6
- Reid, M. J., Webster-Stratton, C., & Beauchaine, T. P. (2001). Parent training in head start: A comparison of program response among African American, Asian American, Caucasian, and Hispanic mothers. Prevention Science: the Official Journal of the Society for Prevention Research, 2, 209-227.
- Reinke, W. M., Herman, K. C., & Dong, N. (2016). The Incredible Years Teacher Classroom Management program: Outcomes from a group randomized trial. Manuscript submitted for publication. Retrieved from http://www. incredibleyears.com/wp-content/uploads/Reinke-IY-TCM-Program-Outcomes.pdf
- Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: Follow-up study of antisocial children into adulthood. BMJ Clinical Research, 323, 191. doi.org/10.1136/bmj.323.7306.191
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin.
- Snell, T., Knapp, M., Healey, A., Guglani, S., Evans-Lacko, S., Fernandez, J.-L., ... Ford, T. (2013). Economic impact of childhood psychiatric disorder on public sector services in Britain: Estimates from national survey data. Journal of Child Psychology and Psychiatry, 54, 977–985. doi.org/10.1111/jcpp.12055
- Snyder, J., Low, S., Schultz, T., Barner, S., Moreno, D., Garst, M., ... Schrepferman, L. (2011). The impact of brief teacher training on classroom management and child behavior in at-risk preschool settings: Mediators and treatment utility. Journal of Applied Developmental Psychology, 32, 336-345. doi.org/10.1016/j.appdev.2011.06.001
- Sørlie, M. A., & Ogden, T. (2015). School-wide positive behavior support-Norway: Impacts on problem behavior and classroom climate. International Journal of School & Educational Psychology, 3, 202-217. doi.org/10.1080/ 21683603.2015.1060912
- Statistics Norway. (2017). Education statistics. Primary and lower secondary schools. Retrieved from https://www.ssb. no/utdanning
- Stuart, E. A., Azur, M., Frangakis, C., & Leaf, P. (2009). Multiple imputation with large data sets: A case study of the children's mental health initiative. American Journal of Epidemiology, 169, 1133-1139. doi.org/10.1093/aje/kwp026
- Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools: What does the evidence say? Health Promotion International, 26, i29-i69. doi.org/10.1093/heapro/dar075
- Webster-Stratton, C. (1999). How to promote children's social and emotional competence. London: Paul Chapman.
- Webster-Stratton, C. (2011). Promoting positive academic and social behaviors: The teachers and children series. The teachers and children series. A video-based course (Lederveiledning - Norwegian ed.). Seattle, WA: Incredible Years.
- Webster-Stratton, C. (2012). Incredible teachers: Nurturing children's social, emotional, and academic competence. Seattle, WA: Incredible Years.
- 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 & Adolescent Psychology, 30, 283-302. 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 & Adolescent Psychology, 33, 105-124. 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, and Allied Disciplines, 49, 471-488. doi.org/10.1111/j.1469-7610.2007.01861.x
- Wyrick, A. J., & Rudasill, K. M. (2009). Parent involvement as a predictor of teacher-child relationship quality in third grade. Early Education & Development, 20, 845-864. doi.org/10.1080/10409280802582803