Journal of School Psychology 86 (2021) 133–150

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

A school-based parenting program for children with attention-deficit/hyperactivity disorder: Impact on paternal caregivers

Gregory A. Fabiano [a,*], Nicole K. Schatz [a], Kellina Lupas [a], Chanelle Gordon [b], Timothy Hayes [a], Devon Tower [a], Tamara Santos Soto [a], Fiona Macphee [a], William E. Pelham Jr [a], Karen Hulme [b]

[a] Florida International University, United States of America
[b] University at Buffalo, SUNY, United States of America

ARTICLE INFO

Action Editor: Austin Johnson

Keywords:
Parenting
ADHD
School-home support
After-school programming

ABSTRACT

Engaging male caregivers within school settings is a major need within the educational field. Paternal engagement may be particularly important for children with attention-deficit/hyperactivity disorder (ADHD). Children with ADHD have increased risk for a number of poor educational outcomes, which may be attenuated by the benefits of positive male caregiver involvement. The Coaching Our Acting Out Children: Heightening Essential Skills (COACHES) program has been illustrated to be an effective approach for engaging, retaining, and improving the parenting of male caregivers of children with ADHD in clinical settings. The present study reports on the efficacy of the COACHES in Schools program, an adaptation intended for deployment in elementary school settings. Sixty-one male caregivers were randomly assigned to COACHES in Schools or a waitlist control. Results indicated that male caregivers in COACHES in Schools used significantly more praise and less negative talk in a parent-child activity relative to male caregivers in the waitlist control at post-treatment and one-month follow-up. Distal outcomes related to child behavior at home and at school were not significantly different. Implications of the results for future studies and continued efforts to engage male caregivers within school settings are discussed.

Parent involvement in school has become a national goal for educators, school districts, and policy-makers. There are numerous effective approaches for engaging and involving parents in their child's school setting (e.g., Pelham et al., 2016; Pfiffner et al., 2016; Reschly & Christenson, 2019; Sheridan et al., 2019) with research suggesting that positive parent involvement in school-related activities can improve academic performance and academic achievement (e.g., Boonk et al., 2018; Jeynes, 2015; Kim & Hill, 2015). Yet, in spite of these promising approaches and the evidence of effectiveness, there remains a disconnect between parent involvement in the child's educational experience and caregivers' active collaboration with teachers (Hornby & Lafaele, 2011).

This disconnect may be even more exacerbated for children with behavioral challenges. Indeed, if many of the communications between educators and caregivers are related to behavioral or academic concerns, parents may disengage from school involvement

* Corresponding author at: Florida International University, Department of Psychology, Center for Children and Families, 4600 Main Street Suite 101, Buffalo, NY 14226, United States of America.
E-mail address: gfabiano@fiu.edu (G.A. Fabiano).

https://doi.org/10.1016/j.jsp.2021.04.002
Received 28 February 2020; Received in revised form 13 January 2021; Accepted 7 April 2021

(Dishion et al., 2004; Hornby & Lafaele, 2011). For children identified as having ADHD, teachers report a disproportionate number of negative communications with parents regarding behavioral and/or academic concerns (Schatz et al., 2021). Parsing the analysis of the disconnect between parents and schools further, lack of involvement may be most pronounced for children's male caregivers (e.g., fathers). Considering general education settings, 65% of mothers are classified as "highly involved" in the child's school as compared to 33% of fathers (Child Trends, 2002; see Downer, 2007, and U.S. Department of Education, 2000, for expanded discussions); put plainly, half as many male caregivers are currently highly engaged within their child's school relative to female caregivers. This lack of male caregiver involvement may be particularly concerning for children with behavioral challenges as research has illustrated that paternal involvement, in particular, may confer a protective impact on long-term development (e.g., Flouri et al., 2016). Thus, increasing paternal involvement is a key target within the development of school-based interventions aimed at improving social and behavioral functioning of students with behavioral challenges. Improving the reach of parenting interventions to include and impact male caregivers has been a consistent call in the research literature for decades (e.g., Downer, 2007; Fabiano, 2007; Lechowicz et al., 2019; Parent et al., 2017; Phares, 1996; Tiano & McNeil, 2005), but schools continue to need actionable approaches for realizing efforts toward increased engagement.

A promising approach for increasing male caregiver involvement and improving parenting is the *Coaching Our Acting-out Children: Heightening Essential Skills* (COACHES) program, which is a parent-child program that integrates two evidence-based treatments for ADHD, including (a) behavioral parent training (Cunningham et al., 1998), and (b) behavioral peer interventions within recreational activities (Fabiano et al., 2014; Pelham et al., 1998; Pelham & Fabiano, 2008). In this intervention program, during each session male caregivers attend a brief parent training class to learn effective parenting strategies while the children learn sports skills. Then, male caregivers and children join one another for a sports game, with the parents practicing the parenting strategies learned during the activity.

The COACHES program has been evaluated in clinical settings and demonstrated improved parenting behavior and male caregivers' ratings of the children's improvement relative to a waitlist control (Fabiano et al., 2012), with participants in the COACHES program having illustrated increased parent and child attendance and engagement in treatment (i.e., in-home assignment completion following the weekly sessions; arriving on time for sessions) relative to standard parent training groups (Fabiano et al., 2009). Recent modifications of the COACHES approach for preschool-aged students illustrated improved parenting behavior (Caserta et al., 2018; Chacko et al., 2018) following participation, and in this program male and female caregivers attended together. Further, Chacko et al. (2018) also demonstrated improvements in distal outcomes related to children's academic skills in early literacy when male caregivers were encouraged to engage in shared book reading during the parent-child activities.

The logic model (Fig. 1) supporting the use of the COACHES program is consistent with the definition of effective male caregiver engagement explicated in McWayne et al. (2013). McWayne et al. (2013) defined two core aspects of positive and effective male

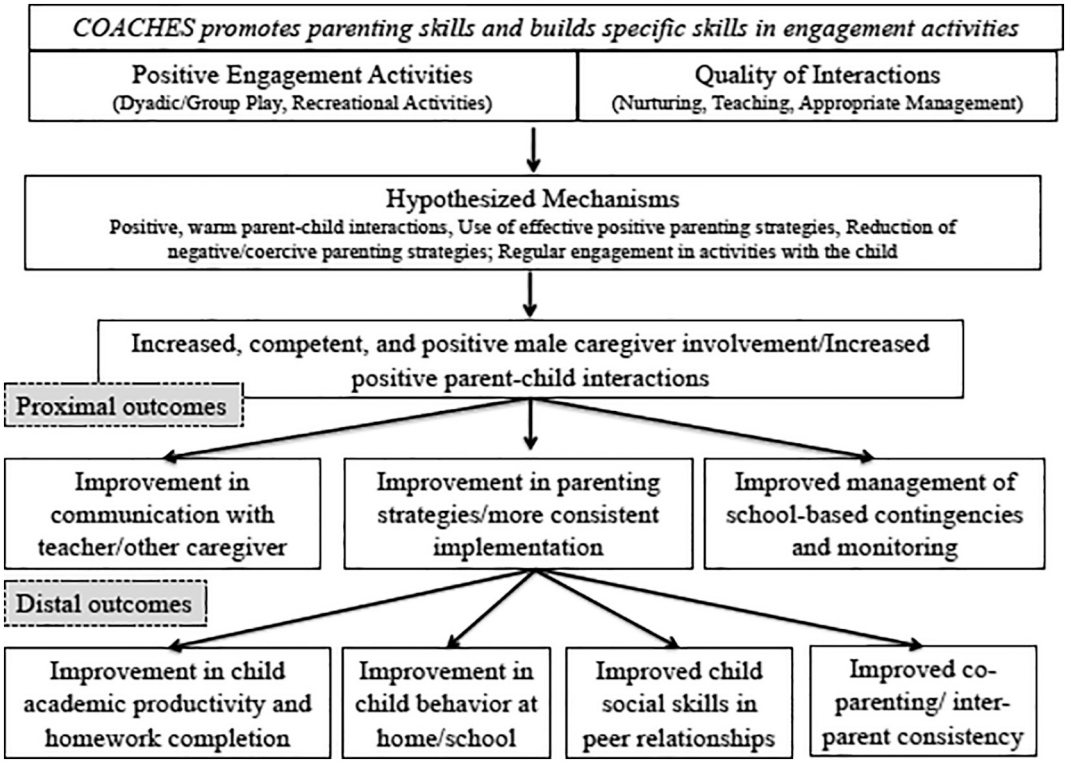

Fig. 1. Logic model illustrating the proposed theory of change in the COACHES in schools program.

caregiver involvement: (a) engaging with the child in positive skill-building activities (distinct from basic childcare tasks such as preparing meals, bathing, etc.), and (b) quality interactions in the parenting role. The COACHES program directly addresses these two domains through the parent-child interactions in authentic activities (e.g., sports) as well as instruction on best practice parenting strategies. COACHES is hypothesized to be an effective approach because it couples engaging and rewarding parent-child activities within a sports context with the content introduced in traditional parenting programs (Fabiano et al., 2009, 2012). Positive engagement activities include the parent-child interactions that occur during the sports activities. Quality interactions are those characterized by liberal praise, use of appropriate commands and limit-setting, and ensuring parent-child engagements are weighted toward more positive (i.e., labeled praise/compliments) and fewer negative (i.e., reprimands, commands, criticisms) interactions.

The logic model for the intervention includes a focus on proximal, parent-focused treatment targets (i.e., parenting skills) that impact distal, child-focused outcomes (e.g., behavioral impairment). The proximal outcomes of the COACHES program include (a) more positive and effective parent-child interactions, and (b) increased use of effective parenting strategies (e.g., effective instructions, contingency management; Webster-Stratton, 2006). These are directly addressed as treatment targets within the COACHES intervention. Distal outcomes reflect improvements in child behavior that are hypothesized to result from the improvements in the parenting approach. These outcomes include increased academic productivity/homework completion, improvement in behavior/deportment at home and at school, and reductions in psychosocial impairment.

The COACHES program has been employed as an intervention to support male caregivers in the parenting of a child with attention-deficit/hyperactivity disorder (ADHD) in clinic settings (Fabiano et al., 2009, 2012). One limitation of the prior work relates to the transportability of the intervention for wide deployment – most clinics do not have the physical space to accommodate group sports activities with multiple families. Furthermore, the exclusive focus on male caregivers as participants in these formative studies marginalized other caregivers who would have reasonable expectations of participating in the intervention as well (e.g., maternal caregivers). Schools provide an alternative location for this approach to ADHD treatment, and alternative treatment settings are highlighted as a mechanism for improving the dissemination and reach of evidence-based treatments (Kazdin & Blasé, 2011; Kazdin & Rabbitt, 2013). A key question regarding the move toward deploying the COACHES program in a community setting relates to whether there would be an impact on participant engagement and attendance. Chacko et al. (2016) reviewed parent training programs that reported attendance, and of those reporting on this variable, parents attended three-quarters of sessions, on average. However, Chacko et al. also noted that only 23% of studies included in the review provided sufficient data on attendance, and the majority of these reports were for maternal caregivers. Parenting programs that also include child activities, such as COACHES, typically result in greater attendance (Chacko et al., 2009; Fabiano et al., 2009; Jensen & Grimes, 2010). Thus, there is an ongoing need to report on attendance for parenting programs, particularly for male caregivers.

The current study aimed to adapt the COACHES program for use within school settings to investigate the feasibility and promise of the approach within a novel context. This included modifying the COACHES program for use in an elementary school and adapting it such that it could be implemented by school personnel. This approach is unique as it is moving the COACHES program from implementation in a clinic to a school setting. Additional adaptations for the school setting included training school personnel to facilitate the implementation of the program and making the program shorter in duration to better fit within the school academic calendar. The primary study question was to investigate whether the COACHES program resulted in improved paternal caregiver (proximal) and child (distal) outcomes relative to a waitlist control condition. It was hypothesized that the COACHES program would result in improved parenting, as measured by observations of parent-child interactions, and that there would be evidence of improvements in distal outcomes, as measured by parent and teacher ratings as well as school-based observations of child behavior. An exploratory aim of the study was to investigate the mediating role of session attendance on improved parenting practices for male caregivers. Whereas attendance has been illustrated to be greater in a COACHES program approach, relative to traditional parenting intervention, no prior study has investigated whether the increased attendance might enhance outcomes. Specifically, it was hypothesized that greater male caregiver attendance would mediate improved parenting.

1. Method

1.1. Participants

Sixty-one male caregivers and their children with ADHD enrolled in the program. Children were in kindergarten through sixth grade. Families were recruited to attend a six-week, after-school program that was free of cost and implemented within elementary schools located in either the Northeastern or Southeastern United States. Each site at the two geographic locations aimed to recruit 30 fathers for the study. Within the Northeastern site ($N = 31$ participants), the program was conducted in two districts. In the first district, the program was conducted at a single school that drew participants from the three elementary schools and a middle school in the district; at the second district the program was held at two elementary schools, each of which included students recruited from another nearby elementary school. At the Southeastern site ($N = 30$ participants) the program was implemented in four schools all from the same district. All program materials were developed in English and Spanish as the majority of participants at the Southeastern site spoke Spanish as the primary language.

Flyers were sent home to families within each school, referrals were made by school professionals (e.g., school counselors, school psychologists, physical education teachers), and informational sessions were held during school open houses. Prospective participants were asked to call a study hotline to complete a phone screening and determine initial eligibility.

Self-identified male caregivers who appeared to meet eligibility criteria based on the phone screening were invited to an in-person meeting where the study was described. Following the discussion, if they chose to participate (a) caregivers signed an informed consent

and provided parental permission for the child to participate, (b) children over the age of seven years provided informed assent, and (c) initial baseline information was collected. Male caregivers, and their children, were the primary focus of this study. Female caregivers and teachers provided diagnostic information at the baseline assessment and on potential generalization of intervention effects at post-intervention and at the one-month follow-up. All procedures were approved by the Institutional Review Boards at each site.

Contemporary diagnostic assessment procedures were used to identify an ADHD diagnosis (American Psychiatric Association, 2013; Pelham et al., 2005). The child's mother, father, and teacher each completed the Disruptive Behavior Disorder (DBD) rating scales of DSM symptoms (Pelham et al., 1992). A semi-structured DBD clinical interview was completed with the child's parents by study clinicians (Hartung et al., 2005; Pelham et al., 2005). Consistent with recommendations in the diagnostic literature, maternal and paternal caregivers completed DBD ratings independently, and ratings were combined across caregivers and the teacher by taking the maximum symptom rating (Anastopoulos & Shelton, 2001; Bird et al., 1992). Cross-situational impairment was assessed through

Table 1
Demographic information for participants in COACHES and waitlist group.

Child	COACHES	Waitlist	Chi-square/t-test
Age in years	8.18 (SD = 1.81)	7.12 (SD =1.67)	t (59) = 2.37, p = .02*
Sex	67.9% Male	78.8% Male	X^2 (1, N = 61) = 0.333, p = .390
Race			X^2 (4, N = 61) = 2.78, p = .596
Asian	3.60%	0.00%	
Black or African American	10.70%	12.10%	
White	75.00%	66.70%	
More than one race	7.10%	9.10%	
Unknown or Other	3.60%	12.10%	
Ethnicity	42.9% Hispanic/Latino	48.5% Hispanic/Latino	X^2 (1, N = 59) = 0.181, p = .795
% taking medication	30.8%	25.0%	X^2 (1, N = 58) = 0.239, p = .769
Child Comorbidity	25% ODD	41.9% ODD	X^2 (1, N = 59) = 1.883, p = .170
Father			
Age in years	43.04 (SD = 9.47)	43.00 (SD = 7.12)	t (49) = 0.013, p = .989
Race			X^2 (4, N = 61) = 0.1.44, p = .837
American Indian/Alaskan Native	0.00%	3.00%	
Asian	3.60%	3.00%	
Black or African American	14.30%	12.10%	
White	67.90%	60.60%	
Unknown or Other	14.30%	21.20%	
Ethnicity	32.1% Hispanic/Latino	42.7% Hispanic/Latino	X^2 (1, N = 54) = 0.458, p = .585
Education			X^2 (6, N = 55) = 3.60, p = .730
Some High School	0.00%	3.00%	
High School or GED	32.10%	30.30%	
Some College	17.90%	12.10%	
Associate's Degree	10.70%	3.00%	
Bachelor's Degree	17.90%	24.20%	
Graduate Degree	7.10%	15.20%	
Marital Status			
Single	7.10%	18.20%	X^2 (4, N = 53) = 3.088, p = .543
Cohabitating with Partner	17.90%	9.10%	
Married	53.60%	48.50%	
Separated	3.60%	3.00%	
Divorced	3.60%	9.10%	
Mother			
Age in Years	38.79 (SD = 7.31)	34.30 (SD = 7.58)	t (49) = 2.15, p = .037*
Race			X^2 (4, N = 61) = 4.92, p = .296
American Indian/Alaskan Native	0.00%	3.00%	
Black or African American	10.70%	6.10%	
White	78.60%	60.60%	
More than one race	0.00%	3.00%	
Unknown or Other	12.20%	27.30%	
Ethnicity	39.3% Hispanic/Latino	33.3% Hispanic/Latino	X^2 (1, N = 53) = 0.121, p = .785
Education			X^2 (6, N = 53) = 8.091, p = .232
Some High School	0.00%	6.10%	
High School or GED	28.60%	33.30%	
Some College	25.00%	12.10%	
Associate's Degree	3.60%	12.10%	
Bachelor's Degree	10.70%	12.10%	
Graduate Degree	21.40%	6.10%	

Note. SD = Standard Deviation. COACHES=Coaching Our ADHD Children: Heightening Essential Skills. AA = African-American. ODD=Oppositional Defiant Disorder. CD=Conduct Disorder. I/O = Inattentive/Overactive factor. O/D = Oppositional defiant factor. DBD-I = Disruptive Behavior Disorders rating scale, Inattentive factor. DBD-H/I = Disruptive Behavior Disorders rating scale, Hyperactive-Impulsive factor. ODD = Disruptive Behavior Disorders rating scale, Oppositional Defiant Disorder factor. t = t value for independent t-test. X^2 are reported Pearson X^2 values. p values for X^2 represent 2-sided tests.

caregiver and teacher ratings on the Impairment Rating Scale (IRS; Fabiano et al., 2006). Children were diagnosed with ADHD following independent review by two licensed psychologists on the study team if the child met DSM symptom criteria for ADHD at home and school by combining symptom endorsements across all informants (i.e., at least six symptoms of inattentive and/or hyperactive/impulsive behavior), and impairment ratings indicated at least one impairment domain at home *and* at school.

Child inclusion criteria included having a diagnosis of ADHD, being enrolled in Grades K–6, and having at least one male caregiver who was able to participate in the program and attend treatment sessions. Child exclusion criteria included having an IQ less than 70, being diagnosed with psychosis or autism spectrum disorder, or being home-schooled. Demographic information for the participants is summarized in Table 1. Fig. 2 illustrates the participant flow through the screening, randomization, and data collection process.

1.2. Procedures

The COACHES program was implemented within schools embedded within the larger school districts. Families were recruited from each school, and random assignment was stratified within school building across two cohorts and two sites. Paternal caregivers within school were randomly assigned to COACHES or a waitlist group.

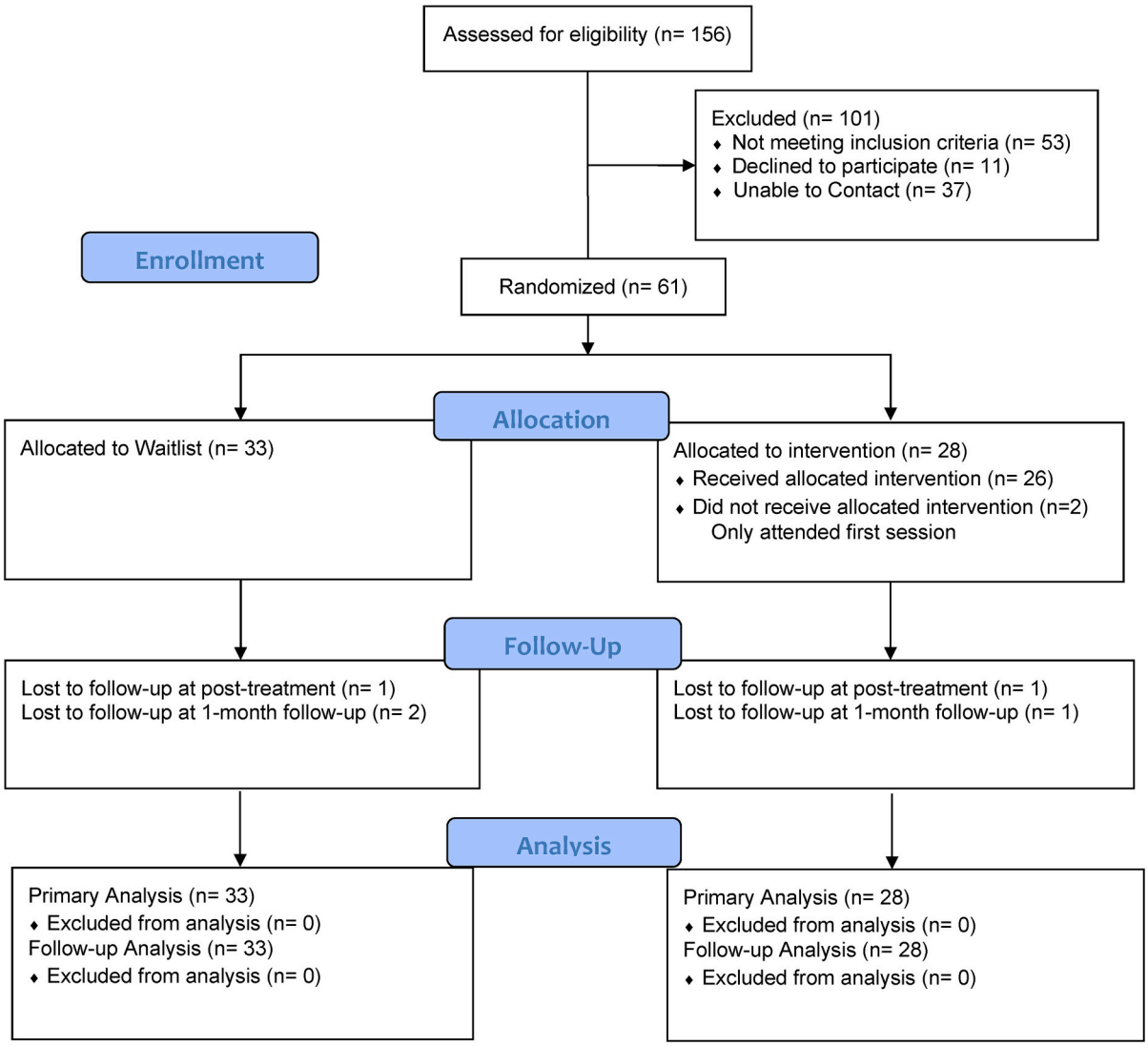

Fig. 2. CONSORT diagram to illustrate study recruitment, screening, random assignment, and data analysis.

1.3. Study groups

1.3.1. COACHES

The COACHES program is based on two evidence-based programs for children with challenging behaviors. The sports skill drills and recreational activities are based on the Summer Treatment Program (Fabiano et al., 2014; Pelham et al., 1998) and the parent training component is based on the Community Parent Education Program (Cunningham et al., 1993, 1995). For the present school-based study, several adaptations to the program were made. Adaptations included (a) a reduction from a 120 min program to a 90 min program and from eight weeks to six weeks; (b) creation of videos to permit the parent training sessions to be implemented without extensive training required for school facilitators; (c) inclusion of all parents and guardians, rather than limiting to solely paternal caregivers to ensure inclusivity (e.g., a child in the school without a male caregiver could still attend; male caregiver's parenting partner[s] could attend); (d) a focus on positive home-school communication and partnership; and (e) the development of all program components (e.g., videos, parent workbook, manual) in English and Spanish. The program session length and number of sessions were reduced due to the constraints of gym availability and to fit the program within school calendars without the interruption of school breaks. Due to the reduction in session length, fewer parenting behaviors could be addressed than in the clinic-based program. Parenting strategies reviewed within the context of the program included (a) an introduction to identifying and modifying the antecedents and consequences of target behaviors and the development of a school-based daily report card, (b) "catching your child being good" by using labeled praise, (c) giving good directions and using effective commands, (d) using Premack contingencies and transitional warnings, (e) organization and planning ahead, and (e) time out.

The COACHES in Schools program was iteratively developed through a series of pilot studies. First, focus groups were conducted with caregivers and school personnel to inform the team on modifications that should be considered for the school setting. The procedures were initially piloted in a set of open trials at both sites, and the investigatory team learned that it would be very difficult to maintain intervention fidelity of the parenting program without extensive training of school professionals and intensive monitoring and supervision. This would be difficult to sustain in school settings and would require repeated intensive training for each school professional who facilitated a group (i.e., if a professional left the school district for another job or could no longer facilitate sessions scheduled in the evening). Therefore, videos of each session were recorded (a) that introduced the skill, (b) that included instructions to organize in subgroups in order to provide opportunities for parents in the group to discuss the skill with other parents, and (c) allowed viewing of multiple exemplars of the skill being used correctly and incorrectly. Consistent with the prior COACHES studies, after viewing the video, parents transitioned to the gym where their children played in a soccer game. During the game, caregivers practiced the skills learned during the meeting and the facilitator discussed strategy use during child water breaks each quarter of the soccer game. At the end of each session, the facilitator elicited suggestions for how caregivers might try the strategy at home and encouraged them to continue practicing the skill throughout the upcoming week. A caregiver workbook was also utilized to facilitate between-session practice. The final week also included a caregiver versus child exhibition game and brief awards ceremony.

Other modifications were also made to the COACHES in Schools program based on school partner and caregiver feedback as well as lessons learned through pilot studies. Because the parents at one site were more likely to speak Spanish, all program materials were created with both an English and Spanish version. The child activity facilitators (typically physical education teachers) managed the child soccer skills period and refereed and organized the child game. Caregiver meeting facilitators (typically school counselors or special education teachers) managed the group video sessions, as well as conducted the brief discussion groups between each game quarter to emphasize skill implementation and problem-solving. Training for both the child activity facilitators and the caregiver group facilitators involved self-review of a training manual. COACHES in Schools topics were modified to focus on outcomes relevant to school settings including a discussion of home-school partnership through the use of a daily report card (Pyle & Fabiano, 2017; Volpe & Fabiano, 2013), organization and planning related to homework (e.g., Langberg, 2011; Schultz & Evans, 2015), and communication with the teacher about the child's behavior and academic performance. Finally, the school partners requested opening the COACHES in Schools program to all caregivers rather than restricting it to solely male caregivers as was done in the clinic-based version (e.g., Fabiano et al., 2009, 2012) in order to be inclusive for all families in the school.

After the development process for the COACHES in Schools program was completed, families randomly assigned to COACHES in Schools participated in the six-week program. After one month, a follow-up evaluation was conducted. The final program implemented included the components and content as outlined above. The program was wholly implemented by internal school staff (e.g., school counselors, physical education teachers, special education teachers) and the research staff conducted treatment fidelity observations and supervision/feedback for the school staff.

1.3.2. Waitlist

Families assigned to the waitlist group were evaluated six weeks after the initiation of the study program and then again, one month later. Following the one-month follow-up assessment, they were then enrolled in the COACHES in Schools program.

2. Measures

The measurement plan focused on the hypothesized proximal (i.e., parenting) outcomes, hypothesized distal (i.e., child behavior) outcomes, and generalization of intervention effects to non-targeted individuals (i.e., another caregiver [typically the mother], the teacher) and settings (e.g., the classroom). The hypothesized mechanism in improved child behavior was improvement in parenting behaviors. For this reason, a behavioral observation of paternal caregiver-child interactions was conducted and coded before intervention, after intervention, and at one-month follow-up.

2.1. Proximal observation of parenting

Male caregiver-child interactions represent the proximal outcome of the parenting program and observations of these interactions are the most common approach for assessing parenting skill use in studies of parent management training. Parenting skills were measured (a) at baseline, (b) after the initial COACHES program ended, and (c) at one-month follow-up for male caregivers in both study groups using the Dyadic Parent-Child Interaction Coding System-III (DPICS-III; Eyberg et al., 2006; Eyberg & Robinson, 1983; Robinson & Eyberg, 1981). The DPICS-III is an observational system used to measure child behavior (e.g., noncompliance) and parenting behavior (e.g., critical statements, praise). The DPICS-III has been used as an outcome in multiple studies of parent training and has demonstrated sensitivity to measuring outcomes for male caregivers (Fabiano et al., 2012; Schuhmann et al., 1998). For the current study, the observation scenario was modified to make it appropriate for the school setting and feasible for parents who were not attending clinic-based sessions. This approach was successfully used in similar work (Caserta et al., 2018). Specifically, each male caregiver was individually asked to help the child complete a soccer skill drill of dribbling around five pylons and shooting on a goal in a private location in the school or at their home in the driveway/yard. The parent stood between two pylons that served as a goal, and the child faced the parent at the end of the row of pylons. A research assistant stood behind the child, facing the parent, and video-recorded the interaction for later coding. The parent and child were given 2 min to complete the task.

Dependent measures used in the study were frequency counts of behaviors from the DPICS-III Manual (Eyberg et al., 2006) and included (a) praise (labeled plus unlabeled praise: "Labeled praise provides a positive evaluation of a specific behavior, activity, or product of the child" [p. 46] and "An Unlabeled Praise provides a positive evaluation of the child, an attribute of the child, or a nonspecific activity, behavior, or product of the child" [p. 53]); (b) commands (direct plus indirect commands: "Direct commands are declarative statements that contain an order or direction for a vocal or motor behavior to be performed and indicate that the child is to perform this behavior" [p. 33]; "An Indirect Command is a suggestion for a vocal or motor behavior to be performed that is implied or stated in question form" [p. 38]); and (c) negative talk ("A verbal expression of disapproval of the child or the child's attributes, activities, products, or choices. Negative talk also includes sassy, sarcastic, rude, or impudent speech" [p. 23]).

Research assistants were trained as coders for the videotaped interactions using the DPICS-III manual and were required to meet 100% accuracy on an operational definitions test and to accurately code sample videos prior to study participant videos. Coders were unaware of treatment condition. Approximately half of the videos (51%) were coded twice to allow the calculation of reliability among coders. Reliability analyses using two-way mixed, consistency, average-measures intraclass correlations (ICCs) for total count of behaviors revealed excellent inter-rater reliability among the study outcome measures of praise (ICC = 0.89), commands (ICC = 0.88), and negative talk (ICC = 0.81).

2.2. Psychiatric symptoms

Attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), and conduct disorder (CD) *Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition* (DSM-V; American Psychiatric Association, 2013) symptoms were measured using the DBD rating scale (Pelham et al., 1992) administered to the caregivers and teachers. At each assessment point, the caregiver was asked to "select the descriptor that best describes the child." The DBD is a 45-item measure that asks parents to rate the DSM symptoms of ADHD, ODD, and CD on a 0–3 point Likert scale (i.e., *Not at all, Just a little, Pretty Much,* or *Very* Much). The 0–3 ratings were summed across 18 ADHD symptoms and then an average score for ratings across symptoms was computed for each individual. For ADHD symptoms, the teacher DBD rating scale has good internal consistency (coefficient alpha = 0.91–0.96, see Pelham et al., 2005; current study coefficient alpha was 0.87–0.94). Across a number of studies, the DBD has also shown sensitivity to behavioral treatment effects (Pelham et al., 2005). Scores on the total ADHD symptom factor could range from 0.0–3.0 and represent the average rating from the ADHD items.

ODD frequently is comorbid with ADHD (American Psychiatric Association, 2013) and often targeted in parenting programs (Kaminski et al., 2008). Thus, ODD symptoms were measured using the DBD rating scale (Pelham et al., 1992). The DBD rating scale factor for ODD symptoms has good internal consistency (coefficient alpha = 0.91–0.96, see Pelham et al., 2005; current study coefficient alpha was 0.90–0.92). Across a number of studies, the DBD has also shown sensitivity to behavioral treatment effects (Pelham et al., 2005). Scores on the total ODD symptom factor could range from 0.0–3.0 and represent the average rating for the ODD items.

2.2.1. Functional impairment

The Impairment Rating Scale (IRS; Evans et al., 2005; Fabiano et al., 2006) asks paternal and maternal caregivers and teachers to rate the severity of the child's problems and need for treatment and/or special services in important functional domains (i.e., relationship with peers, relationship with the caregiver[s]/teacher[s], academic progress, family/classroom functioning, self-esteem, overall need). There are six items on the scale, and scores on the measure range from 0 (*Not a problem/Definitely does not need treatment or special services*) to 6 (*Extreme problem/Definitely needs treatment or special services*). Test-retest reliability ranged from 0.60–0.89 over a period of six months and 0.54–0.76 over one year. Higher ratings on the IRS predict greater likelihood of use of treatment, and there is evidence of convergent and discriminant validity on the measure (Fabiano et al., 2006). The average score for each of the individual domains rated were used in the analysis; scores could range from 0.0–6.0.

2.2.2. Homework impairment

The Homework Problems Checklist (HPC; Anesko et al., 1987; Langberg et al., 2010; Power et al., 2006) is a rating scale completed by paternal and maternal caregivers to assess the child's behavior surrounding homework. The 20-item scale asks caregivers to rate

behaviors as occurring "*Never*" (0) to "*Very Often*" (3). It is comprised of two distinct factors: (a) Factor 1 is labeled "Inattention/Avoidance of Homework", and (b) Factor 2 is labeled "Poor Productivity/Nonadherence with Homework Rules" (Langberg et al., 2010; Power et al., 2006). The HPC demonstrates acceptable levels of internal consistency (Anesko et al., 1987; Langberg et al., 2010); for this study alphas ranged from 0.92–0.95 for Factor I and 0.86–0.88 for Factor II. The HPC factors evidence concurrent validity with other measures of child behavior (e.g., Power et al., 2006). Scores could range from 0 (*Never*) to 3 (*Very often*) and for the present study the average score on each factor was used.

2.2.3. Distal observations of child classroom functioning

Independent observations of the child's classroom behavior were conducted using the Student-Behavior Teacher-Response observation system (SBTR; Pelham et al., 2008) as a measure of generalization. For the present study, the SBTR was used to record the frequency of student rule violations. The SBTR evidences good inter-rater reliability with a correlation between two observers of 0.87 ($p < .001$; Vujnovic et al., 2014). SBTR observations occurred on three separate days for 30 min each. Forty-four of the SBTR observations (25%) were completed simultaneously with a second observer to allow for the calculation of reliability. In this study, inter-rater reliability was assessed for total rule violations using a two-way mixed, consistency, average-measures intra-class correlation, and demonstrated that raters were highly reliable, $ICC = 0.98$.

2.2.4. Social validity

As an indicator of satisfaction at the end of the program, male caregivers were asked to answer the following questions on a four-point scale (e.g., 1 = *Strongly Disagree* to 4 = *Strongly Agree*): (a) Would you attend the program if you could do it over again? (b) Would you recommend the program to other fathers? (c) Please rate your satisfaction with the program compared to other treatment services your child has received, and (d) Please rate how effective the program was in changing your child's problems compared to other treatment services you have received.

3. Analysis plan

Data analyses were completed using R software (R Core Team, 2013). Fig. 3 displays the latent growth curve model used to examine our primary (proximal) and secondary outcome variables. The figure depicts a latent growth curve model, estimated in the structural equation modeling framework, with a treatment dummy variable (0 = waitlist, 1 = COACHES) predicting both the latent intercept and latent slope factors, *I* and *S*, respectively. Note that because an initial set of analyses resulted in warning messages and estimation issues when freeing the latent slope variance and intercept-slope covariance, we simplified all models to only include a random intercept but not a random slope. Thus, Fig. 3 depicts a standard growth model with a random intercept but a fixed slope – a standard multilevel model.

The model in Fig. 3 reflects our hypotheses that treatment condition (i.e., waitlist versus COACHES) would influence both participants' outcomes at end-of-treatment (EOT) and follow-up and, potentially, alter their trajectories over time on our key outcome variables of interest. To test our predictions at both EOT and follow-up, we centered the intercept in two different ways in the models. First, we centered the intercept at EOT by setting the slope factor loadings to [−3 0 + 2].[1] Next, we centered the intercept at follow-up by setting the slope factor loadings to [−5–2 0]. Thus, the effect of treatment on the latent intercept in these two analyses reflected the influence of treatment on participants' scores at EOT and follow-up, respectively.

Fig. 4 displays a path model for our exploratory aim: predicting labeled praise at EOT and Follow-up from total sessions attended by male caregivers. This measure was only collected for participants in the COACHES group; consequently, these analyses reflect only participants in the COACHES condition.

Before proceeding, we note that Structural Equation Models employ Maximum Likelihood (ML) estimation by default, which relies on asymptotic theory and presumes large samples. One may wonder about the effects of these methods in our finite sample of $N = 61$ individuals. In the context of growth models, the multilevel modeling framework offers the possibility of estimating models using residual (restricted) ML (REML, see McNeish, 2017, for an overview) instead of conventional ML. This method provides more precise estimates of model variance components, analogous to dividing variance and covariance terms by $N - 1$ rather than N (see McNeish, 2017). As noted below under the Handling of Missing Data, we ran these same analyses in the multilevel modeling framework with REML estimation and observed only minor numerical difference from the ML results reported here. Similarly, the pattern of significance in the path model of Fig. 4 does not change when adjusting the standard errors using the Ordinary Least Squares denominator correction of $N - p - 1$ (where p = the number of predictors in a given regression) and testing for significance using an appropriate t rather than z-distribution.

Statistical power in growth curve models is a complex function of many parameters (see Brandmaier et al., 2018; Curran et al., 2010) that must be correctly specified to obtain accurate estimates of required sample size (typically using Monte Carlo simulation methods). As such, the exploratory nature of the present research lends a large degree of uncertainty to such an approach. For more concrete rules of thumb, Curran et al. (2010) noted that:

[1] We used these time codings for the slope factor loadings to reflect the unequal spacing of our measurement occasions, with 6 weeks (three two-week increments) elapsing between baseline and EOT and 4 weeks (two, 2-week increments) occurring between EOT and follow-up.

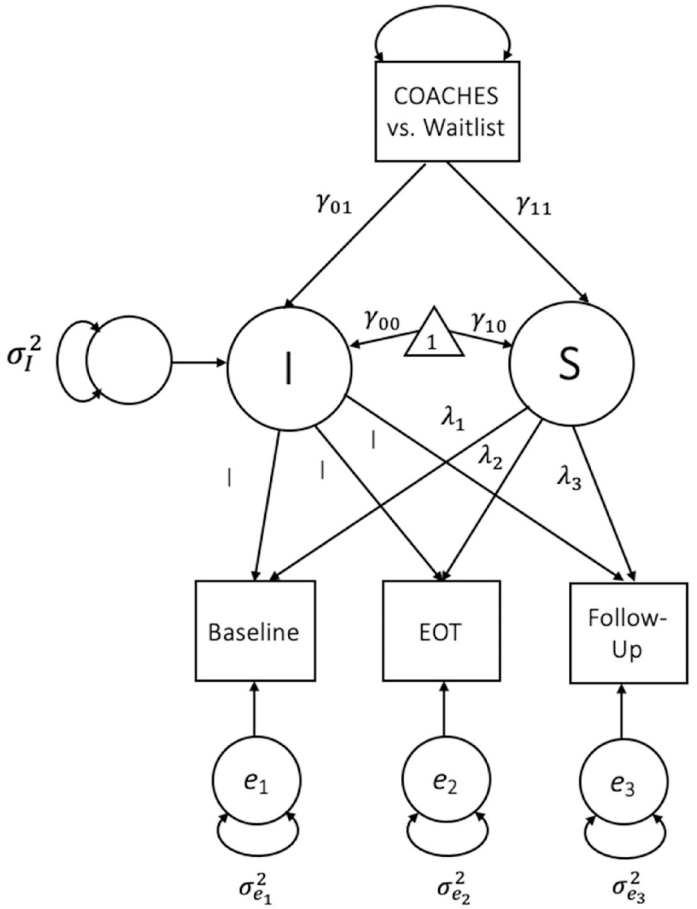

Fig. 3. Random intercept latent growth curve model.
Note. The random intercept was either centered at EOT by setting the slope factor loadings to [−3 0 + 2] or centered at follow-up by setting the factor loadings to [−5–2 0].

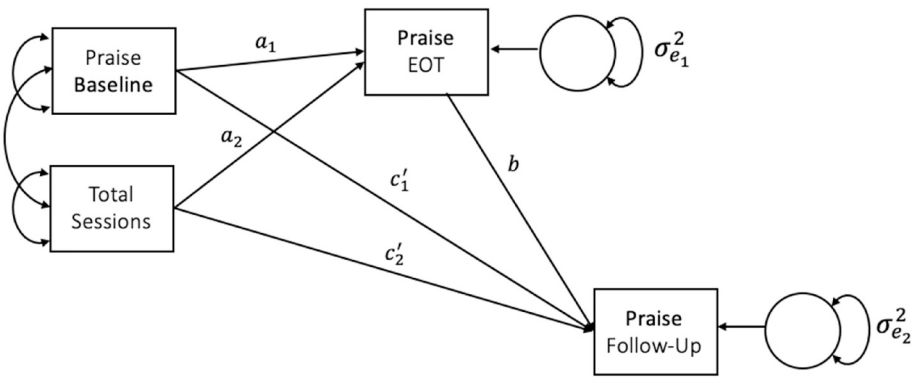

Fig. 4. path model of the exploratory aim, predicting labeled praise at end-of-treatment and follow-up from total sessions attended among male caregivers in the COACHES Group.

> *Growth models have successfully been fitted to samples as small as N = 22 … although sample sizes approaching at least 100 are often preferred. Further, there is a close relation between the number of individuals and the number of repeated observations per individual … as such, the total number of person-by-time observations plays an important role in model estimation and statistical power as well. (p. 125).*

We acknowledge that the sample size of $N = 61$ pilot participants falls between the low $N = 22$ and recommended $N \geq 100$ described by Curran and colleagues. As such, an important goal in our future research is to replicate and extend these findings in a larger sample.

4. Results

4.1. Treatment fidelity and adherence

As noted, the program was manualized and it outlined the skills and procedures to be utilized each week. Treatment fidelity was assessed by having a member of the research team conduct at least one observation of the facilitator over the course of the six weeks. Each session had a checklist of required content that could be reviewed as included or not (e.g., subgroup discussions following video review; parent discussions following each quarter of the sports game to review parenting strategy use). A review of the fidelity forms completed at each site indicated that treatment was implemented as described in the manual. Fidelity was evaluated during 67% of the program sessions and 93% of the session content listed on each meeting checklist was rated by the observer as implemented.

There was variability in attendance to the weekly programs for male caregivers who were consented and enrolled in the COACHES in Schools program. The mean number of sessions attended by male caregivers was 3.2 ($SD = 2.35$) out of six and an additional co-parent or other caregiver attended 3.1 ($SD = 1.83$) sessions. The mean number of sessions attended by children of participating fathers was 3.8 ($SD = 2.24$). Child attendance was greater than caregiver attendance because children could attend with either caregiver present (i.e., both caregivers did not need to be present). Participant data were entered as total sessions attended; thus, it was possible to analyze the total number of sessions that a male caregiver attended, but it was not possible to ascertain the number of sessions attended by caregivers jointly with a partner or individually.

4.2. Handling of missing data

Table 2 displays the proportion of data present for labeled praise and negative talk at each time point as well as the pairwise proportion of cases present (covariance coverage) between measurement occasions on each variable. To handle missing data, we estimated all multilevel growth models in the Structural Equation Modeling (SEM) framework using Full Information Maximum Likelihood estimation (FIML; Arbuckle, 1996). FIML uses all available data to compute the log likelihood and estimate the model parameters. FIML estimation rests on the assumption that the missing data are Missing At Random (MAR; Rubin, 1976) – that is, that the distribution of missing scores for study participants is randomly distributed after conditioning on the true causes of missingness. Although the MAR assumption can never be proven, we conducted a thorough set of preliminary analyses searching for plausible *auxiliary variables* (Collins et al., 2001), which are correlates of both missing data and our model variables that may be used to augment the missing data model and help render the MAR assumption more plausible. An initial set of multiple imputation analyses (Rubin, 1987) using all identified auxiliary variables to impute our main model variables returned a comparable pattern of results to those obtained using FIML alone, without auxiliary variables. Therefore, we report the results of our FIML-based SEM analyses here for consistency across analytic aims, with increased confidence that these results were unaltered by the inclusion of the richer missing data model informed by auxiliary variables.

We note that whereas listwise deletion would have removed half of the cases from each analysis (resulting in $N = 31$ for analyses of labeled praise and $N = 30$ for analyses of negative talk), FIML allowed us to estimate our models using data provided by all 61 study participants. We also note that the Fraction of Missing Information (*FMI*, Rubin, 1987) for our key treatment effects ranged from 0.189 to 0.321, corresponding to *Effective Sample Sizes* (N_{Eff}; Savalei & Rhemtulla, 2012) of 49.47 and 41.42 to detect these effects, respectively. Thus, whereas our auxiliary variable and preliminary imputation analyses increased our confidence in the stability of our estimates across analytic strategies, our FIML analyses partially repaired the standard errors associated with key coefficients, allowing us to recover a portion of the statistical power that would have otherwise been lost in a set of corresponding listwise deletion analyses.

Table 2
Covariance coverage for proximal outcomes.

	Proportion of Cases Present		Covariance Coverage		
	Praise	NT	Baseline	EOT	Follow-Up
Baseline	0.95	0.95	–	0.61	0.62
EOT	0.61	0.61	0.61	–	0.49
Follow-Up	0.64	0.62	0.64	0.51	–

Note. NT = "Negative Talk," EOT = "End of Treatment." Lower triangle of covariance coverage matrix indicates proportion of pairwise cases present for labeled praise. Upper triangle of covariance coverage matrix indicates proportion of pairwise cases present for negative talk.

Table 3
Descriptive statistics and analysis results for male caregiver measures of proximal outcomes.

Measure	COACHES Pre-treatment M (SD)	COACHES Post-treatment M (SD)	COACHES Follow-up M (SD)	WL Pre-treatment M (SD)	WL Post-treatment M (SD)	WL Follow-up M (SD)	Effect size (*g*) post-treatment	Effect size (*g*) follow-up
DPICS								
Total Commands	23.62 (12.43)	14.24 (9.51)	16.41 (11.68)	17.41 (11.04)	13.32 (9.49)	14.67 (13.43)	−0.10	−0.14
Total Praise	2.85 (3.42)	5.12 (5.44)	6.37 (4.22)	2.59 (3.10)	3.30 (5.25)	1.81 (4.15)	0.34**	1.09**
Total NT	2.19 (2.29)	1.50 (1.60)	1.28 (1.37)	2.34 (2.22)	3.10 (3.55)	2.44 (2.23)	0.58*	0.63*

Note. COACHES=Coaching Our Acting-Out Children: Heightening Essential Skills. DPICS = Dyadic Parent-Child Interaction System. WL = Waitlist. NT = Negative Talk. Means and standard deviations in the table reflect descriptive statistics using Full Information Maximum Likelihood imputation. M = Mean. SD=Standard Deviation. Effect sizes were calculated as Hedges' g (For post-test and follow-up scores). ** Significant treatment x time interaction, $p < .001$. * Significant main effect of time, $p < .05$.

Table 4
Growth model analyses of labeled praise and negative talk (fathers).

	Labeled Praise						Negative Talk					
	End of Treatment			Follow-Up			End of Treatment			Follow-Up		
	Est	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*	*Est*	*SE*	*p*
Fixed Effects												
Intercept	2.56	0.62	< 0.001	2.53	0.8	0.002	2.56	0.33	< 0.001	2.63	0.4	< 0.001
Time	−0.02	0.15	0.901	−0.02	0.15	0.901	0.04	0.09	0.696	0.04	0.09	0.696
Treatment	2.35	0.89	0.008	3.74	1.12	0.001	−0.95	0.47	0.043	−1.37	0.56	0.013
Interaction	0.69	0.21	0.001	0.69	0.21	0.001	−0.21	0.13	0.108	−0.21	0.13	0.108
Variance Components												
Intercept	6.73	1.84	< 0.001	6.73	1.84	< 0.001	1.80	0.63	0.004	1.80	0.63	0.004
e_1	4.15	1.54	0.007	4.15	1.54	0.007	3.42	0.85	< 0.001	3.42	0.85	< 0.001
e_2	15.82	4.21	< 0.001	15.82	4.21	< 0.001	3.90	1.09	< 0.001	3.90	1.09	< 0.001
e_3	7.99	2.37	0.001	7.99	2.37	0.001	1.50	0.61	0.014	1.50	0.61	0.014

Note. "End of Treatment" and "Follow-Up" indicate the intercept centering for the model. Treatment refers to the effect of the treatment dummy variable predicting the latent intercept. Interaction refers to the treatment x time interaction implied by the prediction of the latent slope factor by the treatment dummy variable.

4.3. Analysis of proximal outcomes

Table 3 reports descriptive statistics for the measures of parenting behaviors collected during observations, and Table 4 presents the results of our growth model analyses of labeled praise and negative talk among male caregivers. For labeled praise, the analysis illustrated two key effects. First, individuals in the COACHES in Schools group engaged in significantly greater average levels of labeled praise than did individuals in the waitlist group, both at EOT and follow-up. Second, treatment significantly predicted participants' slopes over time at both EOT and follow-up. Across both models, the conditional slope over time among male caregivers in the waitlist group (Treatment = 0) was nonsignificant (−0.02). But the treatment x time interaction coefficient indicated that being in the COACHES in Schools group (Treatment = 1) resulted in a 0.69 unit increase in this conditional slope. Thus, male caregivers in the COACHES group exhibited a positive slope of $-0.02 + 0.69 = 0.67$, indicating that these individuals increased their average frequency of praise statements by 0.67 statements between each measurement occasion.

For negative talk, both models returned significant negative coefficients for the Treatment ➔ Intercept pathway (−0.95 and − 1.37, respectively), indicating that participants in the COACHES group engaged in reduced negative talk relative to those in the waitlist group. In these analyses, treatment did not significantly predict the fixed slope (coefficient − 0.21 was not significant in both models). Thus, there was no significant treatment x time interaction.

Finally, we note that in addition to the analyses of labeled praise and negative talk reported above, we also conducted analyses treating male caregiver use of commands as the outcome, but the models did not yield a significant pattern of results for treatment group or the treatment by time interaction.

4.3.1. Model fit results for proximal outcome analyses

Table 5 presents standard model fit indices for our analyses of labeled praise and negative talk. Both models fit extremely well by the majority of standard fit criteria (see Hu & Bentler, 1999). The model chi-square statistic was non-significant for both labeled praise and negative talk, suggesting an absence of significant misfit. The Root Mean Square Error of Approximation (RMSEA; Steiger, 2016) was zero for both models, suggesting zero misfit per degree of freedom in the population, adjusting for sampling variability (estimation error). Similarly, the Standardized Root Mean Residual (SRMR) was less than 0.08 for both outcomes (see Hu & Bentler, 1999), indicating small average correlation residuals and acceptable overall fit.

Table 5
Fit indices from growth models of labeled praise and negative talk.

	Praise	Negative Talk
$\chi^2(df = 4)$	2.10	3.19
p	0.717	0.526
$\chi^2_B(df = 7)$	16.99	5.74
p_B	0.770	0.599
CFI	1.00	1.00
TLI	1.33	−0.12
RMSEA	0.00	0.00
90% CI	[0, 0.14]	[0, 0.17]
p_{close}	0.717	0.526
SRMR	0.05	0.07

Note. The baseline model used in calculating the CFI and TLI was a no-growth model with no treatment effect, following Widaman and Thompson (2003). $\chi^2$and p are the model chi-square and p-value whereas χ^2_B and p_B are the baseline model chi-square and p-value, respectively. CFI = Comparative Fit Index, TLI = Tucker Lewis Index, RMSEA = Root Mean Square Error of Approximation, p_{close} = the p-value from the test of close fit (H_0: RMSEA < 0.05), and SRMR = Standardized Root Mean Residual. Because growth model fit is identical regardless of where the intercept is centered, we present one set of model fit indices for each target outcome.

In contrast to these indices of model misfit, so-called *incremental fit indices* (see Bentler, 1990) quantify the improvement in fit that one's model provides over a baseline model generally defined as the worst-fitting model acceptable in practice. The standard baseline models assumed by SEM software packages is an *independence model*, in which all variables are specified to have means and variances but zero covariances with one another. In the initial runs of our growth models using this program default, both the Comparative Fit Index (CFI; Bentler, 1990) and Tucker Lewis Index (TLI; Tucker & Lewis, 1973) equaled 1.00 for both models, indicating that our models had recovered 100% of the misfit incurred by the independence model.

However, Widaman and Thompson (2003) showed that the independence model is actually not a nested submodel of a standard growth curve model. Instead, these authors suggested that a no growth model (zero slope mean, variance, or covariance with the random intercept) is an appropriate nested baseline model. In line with their recommendations, we computed the CFI and TLI using a no growth model with zero treatment effect as our nested baseline models, using the `baseline.model` argument in the `fitMeasures()` function in `lavaan` (Rosseel, 2012), reporting these values in Table 5.

For both labeled praise and negative talk, the CFI was 1.00, suggesting perfect recovery of all misfit incurred by the no-growth, no-treatment-effect baseline model. The same was true of the TLI for our labeled praise outcome.[2] However, in the case of negative talk, we note that the TLI value is actually small and negative. Examining the baseline model for this outcome, it is easy to see why. Compared to the large baseline chi-square incurred for labeled praise, the baseline model chi-square value for negative talk is small and nonsignificant, indicating that even a no-growth, no-treatment-effect model provides adequate fit to these data. It is possible that the limited scope of the observational activity provided opportunity for less growth and greater resistance to change than labeled praise. In terms of the incremental fit calculations, whereas the CFI defines fit by subtracting the model and baseline degrees of freedom from their respective chi-square values, the TLI is formulated in terms of the ratios of model and baseline chi-square values to their respective degrees of freedom. It turns out that although the subtraction-based CFI calculations suggest a more flattering picture of incremental fit, the ratio-based TLI calculations suggest that our growth model does not improve fit compared to the baseline model. Nonetheless, we emphasize that this result seems clearly due to the fact that both models – our theoretical model, and the no growth baseline model – provide acceptable fit indicated by small chi-square values relative to their degrees of freedom.

4.4. Analysis of distal outcomes

In addition to the proximal outcomes (i.e., male caregiver parenting behaviors) reported above, distal (i.e., secondary) outcomes were analyzed for both caregivers. None of these secondary outcomes yielded a reliable pattern of significant results, so we omit these models here. Table 6 includes descriptive statistics for the distal outcomes, as well as effect size estimates. Additionally, growth model results and corresponding model fit indices are provided in online supplemental Tables O1 and O2, respectively. Treatment effects in these models failed to reach significance, and model fit indices were considerably poorer than the analyses of our primary outcome variables reported in Table 5. For these reasons, we chose not to highlight these results in our main report. However, in the interest of transparency and offsetting potential file drawer problems, we encourage readers interested in these outcomes to consult the online supplemental results.

4.4.1. Ratings of social validity

Male caregivers generally rated the program in positive terms with descriptives for the following items indicating satisfaction (scale ranged from 1 to 4, with scores of four indicating the greatest level of satisfaction): (a) Would you attend the program if you could do it over again? ($M = 3.55$, $SD = 0.77$); (b) Would you recommend the program to other fathers? ($M = 3.68$, $SD = 0.54$); (c) Please rate your

[2] Note that the formula for the CFI ensures that this measure is bounded by 0 and 1 whereas the TLI, although conceptually bounded in the same way, at times drops below zero or exceeds 1 due to the way this index is formulated.

Table 6
Descriptive statistics and analysis results for male and female caregiver measures of distal outcome.

Measure	COACHES Pre-treatment M (SD)	COACHES Post-treatment M (SD)	COACHES Follow-up M (SD)	WL Pre-treatment M (SD)	WL Post-treatment M (SD)	WL Follow-up M (SD)	Effect size (*g*) pre-treatment to post-treatment	Effect size (*g*) pre-treatment to follow-up
Father								
DBD – ADHD	1.23 (0.63)	1.26 (0.76)	1.17 (0.74)	1.62 (0.56)	1.46 (0.67)	1.48 (0.66)	0.28	0.44
DBD – ODD	0.77 (0.65)	0.87 (0.74)	0.72 (0.71)	1.12 (0.7)	1.05 (0.60)	1.03 (0.65)	0.27	0.46
IRS - Average	2.62 (1.54)	2.47 (1.77)	2.49 (1.92)	2.67 (1.44)	2.54 (1.30)	2.61 (1.67)	0.05	0.07
HPC – Factor 1	1.48 (0.81)	1.37 (0.94)	1.39 (0.97)	1.6 (0.85)	1.70 (0.85)	1.47 (0.92)	0.37	0.08
HPC – Factor 2	0.74 (0.69)	0.62 (0.77)	0.68 (0.79)	0.61 (0.59)	0.83 (0.69)	0.66 (0.55)	0.29	0.03
Mother								
DBD – ADHD	1.61 (0.59)	1.48 (0.64)	1.28 (0.67)	1.7 (0.48)	1.58 (0.59)	1.26 (0.57)	0.16	−0.03
DBD – ODD	1.03 (0.84)	0.96 (0.69)	0.86 (0.82)	1.17 (0.7)	1.17 (0.61)	0.87 (0.54)	0.32	0.01
IRS - Average	3.05 (1.23)	2.81 (1.56)	2.73 (1.95)	3.25 (1.29)	3.41 (1.44)	2.84 (1.62)	0.40	0.06
HPC – Factor 1	1.75 (0.76)	1.54 (0.78)	1.63 (0.77)	2 (0.77)	1.81 (0.85)	1.47 (0.90)	0.33	−0.19
HPC – Factor 2	0.88 (0.82)	0.80 (0.69)	0.79 (0.79)	0.75 (0.52)	0.63 (0.42)	0.71 (0.76)	−0.30	−0.10
Teacher								
DBD – ADHD	1.59 (0.7)	1.5 (0.65)	1.48 (0.72)	1.55 (0.79)	1.47 (0.71)	1.30 (0.73)	−0.04	−0.25
DBD – ODD	0.71 (0.65)	0.74 (0.74)	0.89 (0.81)	0.93 (0.92)	1.01 (0.97)	0.81 (0.89)	0.31	−0.09
IRS - Average	3.63 (1.46)	3.57 (1.61)	3.36 (1.6)	3.2 (1.78)	3.07 (1.72)	2.92 (1.65)	−0.30	−0.27
Observer								
SBTR	16.21 (12.5)	10.89 (9.17)	10.02 (9.47)	14.85 (11.52)	15.48 (13.27)	13.37 (9.83)	0.40	0.35

Note. COACHES=Coaching Our Acting-Out Children: Heightening Essential Skills. DBD=Disruptive Behavior Disorder rating scale. ADHD = Attention deficit/hyperactivity disorder. ODD=Oppositional defiant disorder. IRS=Impairment Rating Scale. HPC=Homework Problems Checklist. SBTR = Student Behavior Teacher Response Observation. Means and standard deviations in the table reflect descriptive statistics using Full Information Maximum Likelihood imputation. M = Mean. SD=Standard Deviation. Effect sizes were calculated as Hedges' *g*.

Table 7
Results of path analysis predicting labeled praise at eot and follow-up from total sessions attended among male caregivers.

Outcome predictor	*Est*	*SE*	*p*
Praise EOT			
Intercept	10.98	3.24	0.001
Baseline	0.42	0.13	0.002
Total Sessions	−1.53	0.60	0.012
Resid. Variance	14.99	6.34	0.018
Praise Follow-Up			
Intercept	−2.47	1.85	0.182
Baseline	0.31	0.17	0.065
EOT	0.46	0.13	< 0.001
Total Sessions	1.25	0.36	0.001
Resid. Variance	8.68	2.15	< 0.001

Note. EOT = "End of Treatment", Resid. = residual.

satisfaction with the program compared to other treatment services your child has received ($M = 3.23$, $SD = 0.90$); and (d) Please rate how effective the program was in changing your child's problems compared to other treatment services you have received ($M = 3.13$, $SD = 0.94$).

4.4.2. Analysis of exploratory path model

Table 7 presents the results of our exploratory path analysis model examining the effect of total sessions attended among male caregivers in the COACHES in Schools group on labeled praise at EOT and follow-up, controlling for praise at prior time points. Results of this path model indicate significant, positive effects of total sessions on praise at both EOT and follow-up, controlling for praise at prior occasions. Thus, male caregivers in the COACHES in Schools group who exhibited greater attendance engaged in significantly more labeled praise at both EOT and follow-up. Examining effect sizes, the increment in R^2 due to adding the total sessions variable to the regression predicting EOT was 0.35, a large effect size by Cohen's (1988) standards. Similarly, the increment in R^2 due to adding the total sessions variable to the regression predicting follow-up was 0.21, a moderate-bordering-on-large effect size by Cohen's (1988) standards. Although these effect sizes should be viewed with some caution, based on their calculation from a small sample of participants ($N = 27$ male caregivers from the COACHES group), these exploratory results suggest that the COACHES in Schools program was more effective when attendance was stronger, as would be predicted.

5. Discussion

Increasing the engagement and support provided by parents of children with ADHD for school outcomes is an important area in need of additional development. The present study aimed to increase male caregiver involvement and effective support through a novel treatment approach – the COACHES in Schools program. Male caregivers of children with ADHD were randomly assigned to COACHES in Schools or a waitlist comparison group and evaluated after six weeks and at a one-month follow-up. Results illustrated considerable improvements in positive parenting for male caregivers who attended COACHES in Schools relative to those assigned to the waitlist group. Results of the impact of these improvements in positive parenting appeared to also lead to positive impacts of distal outcomes (i.e., most effect sizes indicated improvement in the COACHES in Schools group relative to the waitlist), although these results did not reach statistical significance. Support for a hypothesized mediator of the effectiveness of COACHES in Schools (i.e., parent adherence to the intervention) was also demonstrated. Each of these results will be discussed in turn, followed by a consideration for how this study informs future work aimed at engaging male caregivers in efforts to enhance their support for a child with ADHD in academic settings.

The present study replicated prior studies that illustrated that the COACHES approach resulted in improved proximal outcomes related to parenting. When the waitlist control group was compared to the COACHES in Schools program, the present study yielded effect sizes of $g = 0.34$ and $g = 0.58$ for praise and negative talk, respectively. These results are comparable to the COACHES study (Fabiano et al., 2012) that used the same masked observation code for father-child interactions and reported effect sizes of $d = 0.54$ and $d = 0.57$ for praise and negative talk, respectively. The attenuated praise results in the COACHES in Schools study may be explained by the reduced intensity of the intervention (six sessions of 90 min apiece in COACHES in Schools compared to eight sessions of 120 min in the clinical COACHES program), or it may be explained by differences in the tasks observed across the two studies. In the COACHES study conducted in a clinical setting, the observation occurred across a 25 min period and included child-directed play, parent-directed play, and clean up, whereas the COACHES in Schools observation was a 2 min observation of the father coaching the child through a soccer skill drill. It may be that praise was less likely during the COACHES in Schools activity due to less total time of engagement and the nature of the task. That negative talk was significantly reduced in the present study at a rate similar to prior studies is a strong indicator of replication, suggesting a robust impact of the COACHES approach is to teach the male caregivers to reduce this contra-indicated parenting behavior. The present study also extends the prior work by illustrating these changes in parenting following the COACHES in Schools program were maintained at a one-month follow-up observation.

The distal outcomes assessed in this study for both male and female caregivers were not significantly impacted by the COACHES in Schools program. However, an inspection of effect sizes generally suggests improvement following the COACHES in Schools program.

The non-significant findings might be explained by multiple reasons, beyond low power to detect differences in this initial pilot study. First, based on pilot work and conversations with school partners, the intensity of the clinical COACHES program was reduced for COACHES in Schools. This may have resulted in attenuated distal effects. Second, the distal outcomes were collected from settings outside of the treatment program, which might make it harder to generalize the treatment effects observed in the parent-child activities that were similar in content to the COACHES in Schools treatment setting. There is evidence that the effects for ADHD behavioral treatments do not appear to generalize strongly to untreated settings (e.g., Daley et al., 2014). Finally, the program was largely focused on improving parenting skills; a more intensive or multi-modal intervention may be required to impact distal outcomes in specific functional domains (e.g., homework; Breaux et al., 2018; Merrill et al., 2017).

Some explanation for the lack of impact on distal outcomes can also be gleaned from the mediational analysis conducted. This study included variability in parent attendance for the COACHES in Schools intervention, and the improvements in praise rates issued by male caregivers was mediated by attendance at a greater number of treatment sessions. Thus, additional studies should focus on better mechanisms for removing treatment barriers for male caregivers' attendance and also on strategies for improving engagement within the treatment program. For example, a potential side effect of opening the program to all family members in the present study was that for some families, parents would alternate attendance rather than attending together. If the program was a typical community little league, this would have little impact on the child's performance within the setting. However, this approach is inadequate in a program that also includes a training component for caregivers, and future studies may have to address expectations for attendance, and consequences of missing sessions, prior to initiation of the program. It is also important to note that the mediational analyses are tentative given the small sample size. Because male caregivers in the waitlist condition were in a waitlist comparison group, rather than an attention control, there was an inability to analyze attendance in a comparison study group. Nonetheless, it is possible that group would moderate the influence of total sessions on praise. That is, if COACHES in Schools is truly more efficacious than business as usual, then individuals who adhere to the COACHES treatment with greater fidelity should engage in more positive behaviors (such as labeled praise), whereas the fidelity to an inefficacious non-treatment (e.g., attention control) should have little bearing on caregivers' behaviors. This is also an area in need of future study.

Exploratory analyses indicated that attendance in the COACHES in Schools program mediated male caregivers' increased use of labeled praise during a parent-child interaction. This is an important result of the study as it suggests that the greater dose of the COACHES in Schools program received, the greater the improvements in positive parenting, which was the targeted mechanism of the intervention. These results can be contrasted with other studies that indicated attendance did not mediate BPT outcomes (Nix et al., 2009; see Nock & Ferriter, 2005, for a review). One reason that attendance in the COACHES in Schools program mediated improvements in parenting may be the unique aspect of the program that includes active practice and implementation of parenting strategies during the recreational activities. This may serve to enhance caregivers' opportunities for initial implementation and subsequent use of the parenting strategy.

The current study detailed the development of the COACHES in Schools program to enhance male caregiver engagement in treatment for children with ADHD; this pilot study supported the viability of the approach. Program changes to accommodate implementation within school settings included reducing the number of sessions, reducing the total time per session, and developing videos of the parent training component to ensure adequate and consistent delivery while minimizing training burden for facilitators. These changes underscore a tension in disseminating child treatments from the clinic to everyday settings wherein the more intensive and comprehensive treatments developed in clinical settings require adaptation for deployment within the community (Atkins et al., 2016). In the current study there were strong, existing resources within the school setting that supported the COACHES in Schools model (e.g., physical gym space; physical education teachers to provide sports education). There were also limitations such as a lack of providers trained in parenting interventions and a culture within the schools that did not routinely engage male caregivers in school-based intervention. These limitations were addressed through the video-taped delivery of the parent training component that has been shown to be effective in other studies (Kling et al., 2010; McGrath et al., 2011; Webster-Stratton, 1990), and the inclusion of all caregivers in the program (in contrast to the clinical COACHES programs that only included male caregivers; Chacko et al., 2018; Fabiano et al., 2009, 2012). These modifications appeared to be feasible.

5.1. Limitations

This study is not without limitations. First, this was a pilot study of the efficacy of the COACHES in Schools program, which was developed using an iterative, development process to adapt the clinical COACHES program. As such, it is the first attempt to evaluate the efficacy of the program, and replication studies with different school partners and investigatory teams are needed. Second, the proximal outcomes were assessed using a context similar to the treatment setting (e.g., coaching the child in a sports skill drill), and future studies should include additional observational contexts (e.g., completion of homework and other caregiver-directed activities). Third, the range of effect sizes on distal outcomes generally favored COACHES in Schools relative to the waitlist group on indicators of effect size, but the sample size for this study was not powered to detect statistical significance for effects of this magnitude. Relatedly, although there were no significant differences for the majority of variables included in Table 1, maternal age and child age were significantly different between groups, with participants slightly older in the COACHES group. Larger studies with more participants will provide a more reliable assessment of the efficacy of COACHES in Schools on distal outcomes. Finally, the demographic characteristics of the participants mirrored that of the school districts that participated in the study. Additional studies with an increasingly diverse set of school districts and family characteristics are needed to investigate the robustness of these findings.

5.2. Summary and future directions

The COACHES in Schools program was fully developed and feasibly implemented in a school setting by authentic school personnel. Male caregivers and their families were effectively recruited for participation. Results indicated that proximal outcomes (i.e., male caregiver parenting behaviors related to use of praise and reduction in negative talk) were significantly improved at post-treatment and follow-up. The results for parent praise were mediated by attendance at the COACHES in Schools program, as was predicted. Distal outcomes (e.g., ratings of child behavior and observation of classroom functioning) generally yielded small to moderate effect sizes that did not reach statistical significance in this pilot trial. Future work should now focus on how to enhance and extend this model of intervention to improve the engagement and contributions of parents, especially male caregivers, in the efforts to support a child with ADHD in academic settings and situations. Additional studies will also need to explore the potential reasons for the modest effect sizes obtained for distal outcomes (e.g., shorter program duration; inconsistent attendance across multiple family members). The present trial of COACHES in Schools adds to the emerging literature supporting the incorporation of male caregivers in treatment (Fabiano & Caserta, 2018), and it also introduces a means of embedding parent training supports within a novel context. Multiple scholars have called for increased male caregiver engagement and involvement in BPT programs (e.g., Chronis et al., 2004; Fabiano, 2007; Fabiano & Caserta, 2018; Tiano & McNeil, 2005), yet this call remains largely unanswered (Parent et al., 2017). Indeed, male caregiver outcomes are often not included in studies of BPT (Fabiano, 2007). These results indicate that male caregivers can benefit from attending BPT, and their attendance within the programming should be emphasized.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150230 to the State University of New York at Buffalo. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We would also like to thank the participation of the school districts, school personnel, and families, who made this study possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsp.2021.04.002.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington D.C.: American Psychiatric Publications.
Anastopoulos, A. D., & Shelton, T. L. (2001). *Assessing attention-deficit/hyperactivity disorder*. New York: Kluwer Academic Publishing.
Anesko, K. M., Schoiock, G., Ramirez, R., & Levine, F. M. (1987). The homework problems checklist: Assessing children's homework difficulties. *Behav. Assess., 9*, 179–185.
Arbuckle, J. N. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates. Inc.
Atkins, M. S., Rusch, D., Mehta, T. G., & Lakind, D. (2016). Future directions for dissemination and implementation science: Aligning ecological theory and public health to close the research to practice gap. *J. Clin. Child Adolesc. Psychol., 45*(2), 215–226.
Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull., 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238.
Bird, H. R., Gould, M. S., & Staghezza, B. M. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *J. Am. Acad. Child Adolesc. Psychiatry, 31*, 78–85.
Boonk, L., Gijselaers, H. J., Ritzen, H., & Brand-Gruwel, S. (2018). A review of the relationship between parental involvement indicators and academic achievement. *Edu. Res. Rev., 24*, 10–30.
Brandmaier, A. M., von Oertzen, T., Ghisletta, P., Lindenberger, U., & Hertzog, C. (2018). Precision, reliability, and effect size of slope variance in latent growth curve models: Implications for statistical power analysis. *Front. Psychol., 9*, 294. https://doi.org/10.3389/fpsyg.2018.00294.
Breaux, R. P., Langberg, J. M., McLeod, B. D., Molitor, S. J., Smith, Z. R., Bourchtein, E., & Green, C. D. (2018). The importance of therapeutic processes in school-based psychosocial treatment of homework problems in adolescents with ADHD. *J. Consult. Clin. Psychol., 86*(5), 427–438.
Caserta, A., Fabiano, G. A., Hulme, K., Pyle, K., Isaacs, L., & Jerome, S. (2018). A wait-list controlled trial of behavioral parent training for fathers of preschool children. *Evid. Based Pract. Child Adolescent Mental Health, 3*, 106–116.
Chacko, A., Fabiano, G. A., Doctoroff, G., & Fortson, B. (2018). Engaging fathers in effective parenting for preschool children using shared book reading: A randomized controlled trial. *J. Clin. Child Adolesc. Psychol., 47*(1), 79–93.
Chacko, A., Jensen, S. A., Lowry, L. S., Cornwell, M., Chimklis, A., Chan, E., … Pulgarin, B. (2016). Engagement in behavioral parent training: Review of the literature and implications for practice. *Clin. Child. Fam. Psychol. Rev., 19*(3), 204–215.
Chacko, A., Wymbs, B. T., Wymbs, F. A., Pelham, W. E., Swanger-Gagne, M. S., Girio, E., … O'Connor, B. (2009). Enhancing traditional behavioral parent training for single mothers of children with ADHD. *J. Clin. Child Adolesc. Psychol., 38*(2), 206–218.
Child Trends. (2002). *Charting parenthood: A statistical portrait of fathers and mothers in America*. Child Trends: Washington D.C.
Chronis, A. M., Chacko, A., Fabiano, G. A., Wymbs, B. T., & Pelham, W. E. (2004). Enhancements to the behavioral parent training paradigm for families of children with ADHD: Review and future directions. *Clin. Child. Fam. Psychol. Rev., 7*(1), 1–27.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol. Methods, 6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330.
Cunningham, C. E., Bremner, R., & Boyle, M. (1995). Large group community-based parenting programs for families of preschoolers at risk for disruptive behaviour disorders: Utilization, cost effectiveness, and outcome. *J. Child Psychol. Psychiatry Allied Discip., 36*, 1141–1159.
Cunningham, C. E., Bremner, R., & Secord, M. (1998). *The community Parent education program: A school-based family systems oriented workshop for parents of children with disruptive behavior disorders*. Unpublished Treatment Manual.

Cunningham, C. E., Davis, J. R., Bremner, R., Dunn, K. W., & Rzasa, T. (1993). Coping modeling problem solving versus mastery modeling: Effects on adherence, in-session process, and skill acquisition in a residential parent-training program. *J. Consult. Clin. Psychol., 61*, 871–877.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *J. Cogn. Dev., 11*(2), 121–136. https://doi.org/10.1080/15248371003699969.

Daley, D., Van der Oord, S., Ferrin, M., Danckaerts, M., Doepfner, M., Cortese, S., … European ADHD Guidelines Group. (2014). Behavioral interventions in attention-deficit/hyperactivity disorder: A meta-analysis of randomized controlled trials across multiple outcome domains. *J. Am. Acad. Child Adolesc. Psychiatry, 53*(8), 835–847.

Dishion, T. J., Nelson, S. E., & Bullock, B. M. (2004). Premature adolescent autonomy: Parent disengagement and deviant peer process in the amplification of problem behavior. *J. Adolesc., 27*, 515–530.

Downer, J. T. (2007). Father involvement during early childhood. In R. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 329–354). Baltimore: Paul H. Brookes Publishing.

Evans, S. W., Allen, J., Moore, S., & Strauss, V. (2005). Measuring symptoms and functioning in youth with ADHD in middle schools. *J. Abnorm. Child Psychol., 33*, 695–706.

Eyberg, S. M., Nelson, M. M., Duke, M., & Boggs, S. R. (2006). *Manual for the dyadic parent-child interaction coding system* (3rd ed.) Retrieved from: https://www.researchgate.net/publication/242667283_Manual_for_the_Dyadic_Parent-Child_Interaction_Coding_System.

Eyberg, S. M., & Robinson, E. A. (1983). Dyadic parent-child interaction coding system (DPICS): A manual. *Psychol. Doc., 13*, 24.

Fabiano, G. A. (2007). Father participation in behavioral parent training for ADHD: Review and recommendations for increasing inclusion and engagement. *J. Fam. Psychol., 21*, 683–693.

Fabiano, G. A., & Caserta, A. (2018). Future directions in father inclusion, engagement, retention, and positive outcomes in child and adolescent research. *J. Clin. Child Adolesc. Psychol., 47*, 847–862.

Fabiano, G. A., Chacko, A., Pelham, W. E., Robb, J. A., Walker, K. S., Wienke, A. L., Arnold, F., Flammer, L., Keenan, J. K., Visweswaraiah, H., Shulman, S., Herbst, L., & Pirvics, L. (2009). A comparison of behavioral parent training programs for fathers of children with attention-deficit/hyperactivity disorder. *Behav. Ther., 40*, 190–204.

Fabiano, G. A., Pelham, W. E., Cunningham, C. E., Yu, J., Gangloff, B., Buck, M., Linke, S. M., Gormley, M., & Gera, S. (2012). A waitlist-controlled trial of behavioral parent training for fathers of children with attention-deficit/hyperactivity disorder. *J. Clin. Child Adolesc. Psychol., 41*, 337–345.

Fabiano, G. A., Pelham, W. E., Waschbusch, D., Gnagy, E. M., Lahey, B. B., Chronis, A. M., … Burrows-MacLean, L. (2006). A practical impairment measure: Psychometric properties of the impairment rating scale in samples of children with attention-deficit/hyperactivity disorder and two school-based samples. *J. Clin. Child Adolesc. Psychol., 35*, 369–385.

Fabiano, G. A., Schatz, N. K., & Pelham, W. E. (2014). Summer treatment programs for youth with attention-deficit/hyperactivity disorder. *Child Adolesc. Psychiatr. Clin. N. Am., 23*, 757–773.

Flouri, E., Midouhas, E., & Narayanan, M. K. (2016). The relationship between father involvement and child problem behaviour in intact families: A 7-year cross-lagged study. *J. Abnorm. Child Psychol., 44*(5), 1011–1021.

Hartung, C. M., McCarthy, D. M., Milich, R., & Martin, C. A. (2005). Parent-adolescent agreement on ADHD symptoms: A multi-trait, multi-method model. *J. Psychopathol. Behav. Assess., 27*, 159–168.

Hornby, G., & Lafaele, R. (2011). Barriers to parental involvement in education: An explanatory model. *Educ. Rev., 63*(1), 37–52.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J., 6*(1), 1–55. https://doi.org/10.1080/10705519909540118.

Jensen, S. A., & Grimes, L. K. (2010). Increases in parent attendance to behavioral parent training due to concurrent child treatment groups. *Child Youth Care Forum, 39*, 239–251.

Jeynes, W. H. (2015). A meta-analysis: The relationship between father involvement and student academic achievement. *Urban Educ., 50*(4), 387–423.

Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *J. Abnorm. Child Psychol., 36*(4), 567–589.

Kazdin, A. E., & Blasé, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspect. Psychol. Sci., 6*(1), 21–37.

Kazdin, A. E., & Rabbitt, S. M. (2013). Novel models for delivering mental health services and reducing the burdens of mental illness. *Clin. Psychol. Sci., 1*(2), 170–191.

Kim, S. W., & Hill, N. E. (2015). Including fathers in the picture: A meta-analysis of parental involvement and students' academic achievement. *J. Educ. Psychol., 107*(4), 919.

Kling, Å., Forster, M., Sundell, K., & Melin, L. (2010). A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behav. Ther., 41*(4), 530–542.

Langberg, J. M. (2011). *Homework, organization, and planning skills (HOPS) interventions: A treatment manual*. National Association of School Psychologists.

Langberg, J. M., Arnold, L. E., Flowers, A. M., Altaye, M., Epstein, J. M., & Molina, B. S. G. (2010). Assessing homework problems in children with ADHD: Validation of a parent-report measure and evaluation of homework performance patterns. *Sch. Ment. Heal., 2*, 3–12.

Lechowicz, M. E., Jiang, Y., Tully, L. A., Burn, M. T., Collins, D. A., Hawes, D. J., … Frick, P. J. (2019). Enhancing father engagement in parenting programs: Translating research into practice recommendations. *Aust. Psychol., 54*(2), 83–89.

McGrath, P. J., Lingley-Pottie, P., Thurston, C., MacLean, C., Cunningham, C., Waschbusch, D. A., Watters, C., Stewart, S., Bagnell, A., Santor, D., & Chaplin, W. (2011). Telephone-based mental health interventions for child disruptive behavior disorder or anxiety disorders: Randomized trials and overall analysis. *J. Am. Acad. Child Adolesc. Psychiatry, 50*, 1162–1172.

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivar. Behav. Res., 52*(5), 661–670. https://doi.org/10.1080/00273171.2017.1344538.

McWayne, C., Downer, J. T., Campos, R., & Harris, R. D. (2013). Father involvement during early childhood and its association with children's early learning: A meta-analysis. *Early Educ. Dev., 24*(6), 898–922.

Merrill, B. M., Morrow, A. S., Altszuler, A. R., Macphee, F. L., Gnagy, E. M., Greiner, A. R., … Pelham, W. E. (2017). Improving homework performance among children with ADHD: A randomized clinical trial. *J. Consult. Clin. Psychol., 85*(2), 111–122.

Nix, R. L., Bierman, K. L., & McMahon, R. J. (2009). How attendance and quality of participation affect treatment response to parent management training. *J. Consult. Clin. Psychol., 77*(3), 429–438.

Nock, M. K., & Ferriter, C. (2005). Parent management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clin. Child. Fam. Psychol. Rev., 8*(2), 149–166.

Parent, J., Forehand, R., Pomerantz, H., Peisch, V., & Seehuus, M. (2017). Father participation in child psychopathology research. *J. Abnorm. Child Psychol., 45*(7), 1259–1270.

Pelham, W. E., & Fabiano, G. A. (2008). Evidence-based psychosocial treatment for ADHD: An update. *J. Clin. Child Adolesc. Psychol., 37*, 184–214.

Pelham, W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention-deficit/hyperactivity disorder in children and adolescents. *J. Clin. Child Adolesc. Psychol., 34*, 449–476.

Pelham, W. E., Fabiano, G. A., Waxmonsky, J. G., Greiner, A. R., Gnagy, E. M., Pelham, W. E., III, Coxe, S., Verley, J., Bhatia, I., Hart, K., Karch, K., Konijendijk, E., Tresco, K., Nabum-Shani, I., & Murphy, S. A. (2016). Treatment sequencing for childhood ADHD: A multiple-randomization study of adaptive medication and behavioral interventions. *J. Clin. Child Adolesc. Psychol., 45*, 396–415.

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms of the disruptive behavior disorders. *J. Am. Acad. Child Adolesc. Psychiatry, 31*, 210–218.

Pelham, W. E., Greiner, A. R., & Gnagy, E. M. (1998). *The summer treatment program for ADHD*. Unpublished treatment manual.

Pelham, W. E., Greiner, A. R., & Gnagy, E. M. (2008). *Student behavior teacher response observation code manual*. Unpublished observation code manual.

Pfiffner, L. J., Rooney, M., Haack, L., Villodas, M., Delucchi, K., & McBurnett, K. (2016). A randomized controlled trial of a school-implemented school–home intervention for attention-deficit/hyperactivity disorder symptoms and impairment. *J. Am. Acad. Child Adolesc. Psychiatry, 55*(9), 762–770.
Phares, V. (1996). *Fathers and developmental psychopathology*. New York: John Wiley and Sons, Inc.
Power, T. J., Werba, B. E., Watkins, M. W., Angelucci, J. G., & Eiraldi, R. B. (2006). Patterns of parent-reported homework problems among ADHD-referred and non-referred children. *Sch. Psychol. Q., 21*, 13–33.
Pyle, K., & Fabiano, G. A. (2017). A meta-analysis of single-subject design studies utilizing the daily report card intervention for students with ADHD. *Except. Child., 83*, 378–395.
R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://r-project.org/.
Reschly, A. L., & Christenson, S. L. (2019). Family engagement, partnerships, and school support personnel: Connections for enhancing student outcomes. In S. B. Sheldon, & T. A. Turner-Vorbeck (Eds.), *The Wiley handbook of family, school, and community relationships in education* (pp. 203–225). New York, NY: John Wiley and Sons, Inc.
Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent-child interaction coding system: Standardization and validation. *J. Consult. Clin. Psychol., 49*, 245–250.
Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *J. Stat. Softw., 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02.
Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592. https://doi.org/10.2307/2335739.
Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
Savalei, V., & Rhemtulla, M. (2012). On obtaining estimates of the fraction of missing information from full information maximum likelihood. *Struct. Equ. Model. Multidiscip. J., 19*(3), 477–494. https://doi.org/10.1080/10705511.2012.687669.
Schatz, N. K., Fabiano, G. A., Raiker, J. S., Hayes, T., & Pelham, W. E (2021). Twenty-year trends in elementary teachers' beliefs about best practices for students with ADHD. *Sch. Psychol.* In press.
Schuhmann, E. M., Foote, R. C., Eyberg, S. M., Boggs, S. R., & Algina, J. (1998). Efficacy of parent-child interaction therapy: Interim report of a randomized trial with short-term maintenance. *J. Clin. Child Psychol., 27*, 34–45.
Schultz, B. K., & Evans, S. W. (2015). *A practical guide to implementing school-based interventions for adolescents with ADHD*. Springer.
Sheridan, S. M., Smith, T. E., Moorman Kim, E., Beretvas, S. N., & Park, S. (2019). A meta-analysis of family-school interventions and children's social-emotional functioning: Moderators and components of efficacy. *Rev. Educ. Res., 89*(2), 296–332.
Steiger, J. H. (2016). Notes on the Steiger–Lind (1980) handout. *Struct. Equ. Model. Multidiscip. J., 23*(6), 777–781.
Tiano, J. D., & McNeil, C. B. (2005). The inclusion of fathers in behavioral parent training: A critical evaluation. *Child Family Behav. Ther., 27*, 1–28.
Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1–10. https://doi.org/10.1007/BF02291170.
U.S. Department of Education and U.S. Department of Health and Human Services. (2000). *A call to commitment: Fathers' involvement in children's learning*. Washington, D.C.: Author.
Volpe, R., & Fabiano, G. A. (2013). *Daily behavior report cards: An evidence-based system of assessment and intervention*. New York: The Guilford Press.
Vujnovic, R. K., Fabiano, G. A., Pelham, W. E., Greiner, A., Waschbusch, D. A., Gera, S., … Buck, M. (2014). The student behavior teacher response (SBTR) system: Preliminary psychometric properties of an observation system to assess teachers' use of effective behavior management strategies in preschool classrooms. *Educ. Treat. Child., 37*, 323–346.
Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *J. Abnorm. Child Psychol., 18*, 479–492.
Webster-Stratton, C. (2006). *The incredible years: A trouble-shooting guide for parents of children aged 2–8 years* (The Incredible Years).
Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychol. Methods, 8*(1), 16–37. https://doi.org/10.1037/1082-989X.8.1.16.