Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy 50 (2019) 1–14

Behavior Therapy

www.elsevier.com/locate/bt

Effects of Intensive Behavioral Treatment for Children With Varying Levels of Conduct Problems and Callous-Unemotional Traits

Pevitr S. Bansal
Daniel A. Waschbusch
Sarah M. Haas
Dara E. Babinski
Penn State Hershey Medical Center

Sara King
Mount St. Vincent University

Brendan F. Andrade
University of Toronto

Michael T. Willoughby
RTI International

The purpose of this study was to examine whether callous–unemotional (CU) traits moderated the effects of intensive behavior therapy in elementary school-age children with varying levels of conduct problems (CP). Both treatment response (magnitude of change between pre- and post-treatment) and treatment outcomes (likelihood of normalization from treatment) were examined. Participants were 67 children (n = 49 boys, M_{age} = 9.6 years) with varying levels of CP and CU who participated in an intensive 8-week summer treatment program (STP) in which behavior therapy was delivered to children in recreational and classroom settings and to parents via weekly parent training sessions. Effects of treatment were measured using parent and teacher ratings of oppositional defiant disorder (ODD), conduct disorder (CD), callous behavior, and impairment. Results showed that CU moderated treatment effects for CD and callous behavior but not ODD or impairment. The moderating effects showed some evidence that participants with high CP and high CU before treatment had better treatment responses (larger change between pre- and posttreatment) but worse treatment outcomes (lower likelihood of normalization after treatment). These results suggest that intensive treatment, such as the STP, may be necessary but not sufficient for children with CP and CU traits.

Keywords: conduct problems; callous–unemotional; behavior therapy; treatment response; treatment outcomes

We would like to thank the many children, parents, teachers, counselors, and research assistants who made this project possible.

This project was supported by grants to D.A.W. from the Nova Scotia Health Research Foundation (304e) and the Social Sciences and Humanities Research Council of Canada (Grants 839-2000-1061, 410-2004-1272).

Address correspondence to Daniel A. Waschbusch, Ph.D., Penn State Hershey Medical Center, Department of Psychiatry H073, 500 University Drive, Hershey, PA 17033; e-mail: dwaschbusch@pennstatehealth.psu.edu.

DECADES OF RESEARCH HAVE DEMONSTRATED that behavioral treatment (BT) is an effective, empirically supported intervention for oppositional defiant disorder (ODD) and conduct disorder (CD) in youth

(Brestan & Eyberg, 1998; Eyberg, Nelson, & Boggs, 2008) with direct comparisons showing that BT is more effective than non-BT (Comer, Chow, Chan, Cooper-Vince, & Wilson, 2013; Stattin, Enebrink, Özdemir, & Giannotta, 2015). Yet even the most successful interventions are not effective for everyone, and research suggests that roughly one third of children with conduct problems (CP) show insufficient improvement from BT (Kolko et al., 2009; Ollendick et al., 2015; Reid, Webster-Stratton, & Hammond, 2003). Identifying which children do not sufficiently respond to BT is important because it may lead to better approaches to helping them.

One promising indicator in this regard is callous–unemotional (CU) traits. Children with CU traits display an interpersonal–affective style characterized by a lack of empathy or guilt, shallow and deficient emotions, and lack of concern about personal performance (Frick, Ray, Thornton, & Kahn, 2014). Two independent reviews recently concluded that in children with CP, higher CU traits are often associated with less benefit from BT (Frick et al., 2014; Hawes, Price, & Dadds, 2014). At the same time, not all studies support this conclusion, with some studies showing positive effects, especially on measures of CU (Hyde, Shaw, Gardner, & Cheong, 2013; Kolko & Pardini, 2010). This cross-study variability in treatment effects is important to examine because it may provide a clue about how to develop, personalize, and ultimately enhance treatment for children with CU traits. At least three factors that have been suggested as reasons for differences in treatment response across studies: the method of delivering BT, informant differences, and how "benefit from treatment" is defined (Waller, Gardner, & Hyde, 2013; Wilkinson, Waller, & Viding, 2016). These factors are discussed next.

Method of Treatment Delivery

First, the extent to which CU traits influence BT effects may depend on how BT is delivered. A recent review reported that about one half of CU treatment studies delivered BT exclusively to parents (via parent training), with remaining studies delivering BT to both parent and youth (Hawes et al., 2014). Treatments delivered via parent training showed that only higher levels of pretreatment CU traits were consistently associated with a worse treatment response on measures of CP. However, studies of treatments delivered to both parent and youth showed more promise in reducing CP, a finding supported by another review (Wilkinson et al., 2016). These reviews suggest that children with higher levels of CU may benefit from treatments that deliver BT to both the parent (via parent training) and to the youth.

If true, then the summer treatment program (STP) may be a particularly promising treatment for these children. This is because the STP directly targets disruptive behavior by delivering BT directly to children in a summer camp setting and to parents in weekly parent training. The STP has considerable empirical support for children whose primary diagnosis is attention deficit/hyperactivity disorder (ADHD; Pelham et al., 2000)—most of whom also have CP (Waschbusch, 2002)—as well as for children whose primary diagnosis is CP (Kolko, 1995; Kolko, Bukstein, & Barron, 1999). To date, four studies have examined the effects of an STP on children with CU traits (Haas et al., 2011; Miller et al., 2014; Waschbusch et al., 2016; Waschbusch, Carrey, Willoughby, King, & Andrade, 2007). All four studies provided some evidence that BT was less effective for children with high levels of CU traits relative to children with low levels of CU traits. However, these studies examined treatment effects *during* treatment; no research has examined treatment effects using data collected *after* treatment. This is an important point because some studies suggest that children with elevated levels of CU show significant benefit from BT even though they tend to display more disruptive behavior during treatment (Wilkinson et al., 2016).

Informant Differences

A second factor that may be related to differential BT effects across CU treatment studies is informant differences. Research has made it clear that parent and teacher ratings tend to be only modestly correlated (Achenbach, McConaughy, & Howell, 1987), and that different informants provide distinct and meaningful information about children's functioning (De Los Reyes, 2011). Considering these conclusions, it is somewhat surprising that CP treatment studies have generally relied exclusively on parent ratings as the primary treatment outcome measure. For example, one review found that teacher-rated outcomes were used in just 22% of CP treatment studies (Comer et al., 2013), and the same percentage is apparent in treatment studies incorporating CU traits (Hawes et al., 2014). This represents a serious gap because treatment effects may differ across parent and teacher ratings for several reasons. First, parents are often the motivating force behind enrolling children in treatment and almost always actively participate in treatment of CP. Second, CP typically emerge earlier (Loeber, Green, Lahey, Christ, & Frick, 1992) and are more severe (Burns et al., 2013) in home versus school settings. Third, teachers likely have better knowledge of typical child development and have unique information about functioning in the school and

with peers. Importantly, these differences are backed up by a recent meta-analysis of five decades of youth psychotherapy research, which reported that informant was a significant moderator of treatment, with treatment effects about one half as large when reported by teachers as when reported by parents (Weisz et al., 2017). These findings argue for using both parent and teacher ratings when examining the impact of CU on CP treatment, but few studies have done so.

Defining Treatment Response and Outcome

The third factor that may be related to differential BT effects across treatment studies is how treatment effects are defined. Specifically, treatment *response* can be distinguished from treatment *outcome*, with treatment response referring to the magnitude of change between pre- and posttreatment and treatment outcome, referring to whether patients are normalized at the end of treatment (Lindhiem, Kolko, & Cheng, 2012). These different means of measuring treatment effectiveness can lead to different conclusions. For example, severity of CP before treatment has been shown to predict better treatment response (i.e., a larger magnitude of change between pre- and posttreatment) but worse treatment outcome (i.e., more severe antisocial behavior at posttreatment; Reyno & McGrath, 2006). The importance of differentiating these treatment effects is also apparent in CU treatment research—reviews have noted that BT seems to produce improvement but not normalization when delivered to children with elevated CU traits (Hawes et al., 2014; Wilkinson et al., 2016). These findings suggest that children with CP and higher CU traits may have a better treatment response but a worse treatment outcome than children with CP and lower CU. Thus, evaluation of this hypothesis is warranted.

The purpose of this study was to examine whether CU traits moderated the association between CP and effects of BT as delivered in an STP, using parent and teacher ratings to measure treatment response and treatment outcome. It was hypothesized that treatment response and treatment outcome would each be associated with pretreatment CP, but the associations would be moderated by CU traits. However, the nature of the moderating effect was expected to differ for treatment response versus treatment outcome. Specifically, high CP and high CU at pretreatment were expected to be associated with a better treatment response relative to high CP and low CU at pretreatment. In contrast, high CP and high CU at pretreatment were expected to be associated with worse treatment outcomes relative to high CP and low CU at pretreatment. It was also hypothesized that treatment response ratings would be larger for parents than teachers.

Method

PARTICIPANTS

Participants were enrolled in an STP in Atlantic Canada from 2001 to 2002. The sample included 67 children (grade: $M = 3.4$, $SD = 1.7$), including 49 boys and 18 girls, with a mean age of 9.6 ($SD = 1.7$) at pretreatment. The average IQ, estimated from the block design and vocabulary subtests of the Wechsler Intelligence Scale for Children—3rd edition (Wechsler, 1991), was 97.8 ($SD = 16.3$). Most of the sample ($n = 44$, 78.6%) identified as Caucasian, and the average family income (in 2003 Canadian dollars) was \$40,000 ($SD = \$29{,}000$). Exclusion criteria were age younger than 7 or older than 12.99 on the first day of treatment; estimated full-scale IQ < 70; past or current diagnosis of pervasive developmental disorders, schizophrenia or other psychotic disorders, sexual disorders, organic mental disorders, or eating disorders; physical or medical condition that would preclude participation in the STP; and psychopharmacological treatment other than stimulant medication. Most participants ($n = 46$, 68.7%) met criteria for ADHD, ODD, and/or CD, while 21 children (31.3%) did not meet criteria for ADHD, ODD, or CD and participated solely for research purposes (controls). The number of participants (% of full sample) in each diagnostic group was Control = 21 (31.3%), ADHD-only = 6 (9.0%), ODD-only = 2 (3.0%), ODD/ADHD = 22 (32.8%), CD/ADHD = 1 (1.4%), and ODD/CD/ADHD = 15 (22.4%). Of those who met diagnostic criteria for ADHD and/or CP, about one half were medicated during treatment ($n = 24$, 52.2% of diagnosed sample). Prior to treatment, scores on the parent–teacher combined CU scale from the Antisocial Process Screening Device (Frick & Hare, 2001) ranged from 0 to 12 ($M = 6.21$, $SD = 2.77$). *T* scores from the CU scale ranged from 37 to 78 ($M = 58.09$, $SD = 10.91$) and 18 of 67 participants (26.9%) had CU *t* scores above 65, including 14 with ODD or CD and 4 without ODD or CD.

PROCEDURES

The study was approved by a university-based Institutional Review Board and written informed consent was obtained from parents and verbal assent from children. Families were recruited by providing parents information about the treatment study through fliers sent to schools and mental health professionals, advertisements on radio and TV, and postings in public locations (e.g., libraries, grocery stores, and community centers). Parents who contacted the clinic were given more information and a brief screening evaluation, then (if still interested and eligible) parent and teacher ratings were distributed to be completed and returned.

Ratings were completed an average of 8.60 weeks (*SD* = 6.77) before the first day of treatment (pretreatment) and 11.04 weeks (*SD* = 2.94) after the last day of treatment (posttreatment). Parent ratings were available for all participants at pretreatment (n = 67) and for 58 participants at posttreatment, resulting in a parent rating retention rate of 86.6%. Teacher ratings were available for 62 participants at pretreatment (92.5%) and for 47 participants at posttreatment (70.1%), resulting in a teacher rating retention rate of 75.8%. Pretreatment scores did not differ between participants who were retained versus not retained at posttreatment for teacher ratings of CP (p = .545) or CU (p = .876), and parent ratings of CP (p = .284) or CU (p = .205).

Diagnoses were determined using criteria specified in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV; American Psychiatric Association, 1994). Following recommended guidelines (McMahon & Frick, 2005; Pelham, Fabiano, & Massetti, 2005), diagnoses were based on both symptom and impairment criteria as evaluated using multiple sources of information: parent ratings, teacher ratings, and a structured clinical interview administered to parents. Doctoral-level clinicians assigned final diagnoses.

Treatment

The STP is a manualized, evidenced-based program that incorporates multiple aspects of BT into an intensive treatment package (Pelham & Hoza, 1996). The STP was conducted for 8 weeks, Monday through Friday, 8 a.m.–5 p.m. Each day children engaged in 6 hours of structured sports skills (e.g., basketball, soccer, softball, and swimming) and 3 hours of academic learning activities (reading, math, art, and computer). Children were supervised by undergraduate- and graduate-level counselors. Doctoral-level clinicians extensively trained and directly supervised the counselors to ensure accurate and effective implementation of the manualized treatment. Multiple aspects of BT were implemented throughout the STP, including a comprehensive token economy (e.g., "point system"), ubiquitous use of positive reinforcement, a social skills training program, an individualized daily behavior report card, sport skills training and practice, "time out" from positive reinforcement, and parent training in behavior management conducted 2 hours each week in a group format by master's-level clinicians. On average, children attended 96.7% of treatment (38.7 out of 40 days) and parents attended 65% of the parent training (5.2 out of 8 sessions). More detail about the STP is available elsewhere (Pelham et al., 1996, 2010).

MEASURES

Means, standard deviations, and alpha reliabilities for all measures are in Table 1, along with standardized mean difference effect sizes to quantify the change in scores between pre- and posttreatment. At pretreatment, the average correlations between measures were .62 within parent ratings, .65 within teacher ratings, and .46 between parent and teacher ratings. At posttreatment, the average correlations between measures were .58 within parent ratings, .59 within teacher ratings, and .38 between parent and teacher ratings. A table presenting all correlations between measures, separately by informant and treatment, is available as an online supplement.

Treatment Predictors/Moderators

Disruptive Behavior Disorders Rating Scale (DBDRS). The DBDRS consists of 45 items that include the DSM-IV symptoms of ADHD, ODD, and CD (Pelham, Gnagy, Greenslade, & Milich, 1992). Items are rated using Likert scales that range from 0 (*not at all*) to 3 (*very much*). ADHD-Inattentive and ADHD-Hyperactive/Impulsive items were averaged to compute an ADHD severity score and ODD and CD items were averaged to compute a CP severity score.

Antisocial Process Screening Device (APSD). The APSD consists of 20 items, including six items that measure CU traits (Frick & Hare, 2001). Items are rated using Likert scales ranging from 0 (*not at all true*) to 2 (*definitely true*). Positively phrased items on the CU scale were reverse scored, summed (parent pretreatment: *M* = 4.39, *SD* = 2.53; parent posttreatment: *M* = 3.50, *SD* = 2.24; teacher pretreatment: *M* = 4.85, *SD* = 2.97; teacher posttreatment: *M* = 3.80, *SD* = 2.96), and then converted into t scores using published norms.

Treatment Outcomes

Assessment of Disruptive Symptoms DSM-IV (ADS-IV). The ADS-IV (Waschbusch & Sparks, 2003) consists of DSM-IV symptoms of ADHD and ODD, as well as items assessing impairment from symptoms. The symptom items that were used in this study are evaluated by rating how the child compares to others of the same age and sex using Likert scales ranging from 0 (*much less than other children*) to 4 (*much more than other children*), with 2 (*same as other children*) as the midpoint. Items assessing ODD symptoms were averaged to compute an ODD severity score.

Conduct Disorder Rating Scale (CDRS). The CDRS (Waschbusch & Elgar, 2007) consists of DSM-IV symptoms of CD, as well as items assessing impairment from symptoms. The symptom items

Table 1
Sample Sizes (*N*), Means (*M*), Standard Deviations (*SD*), Alpha Reliabilities (α) and Effect Sizes for Measures at Pretreatment and Posttreatment Separately for Parent and Teacher Ratings

	Pretreatment				Posttreatment				SMD
Parent ratings	*N*	*M*	*SD*	α	*N*	*M*	*SD*	α	effect size
1. DBD CP	67	0.52	0.40	.91	58	0.40	0.31	.87	0.23
2. APSD CU	67	59.79	12.40	.72	58	55.45	10.83	.62	0.37
3. ADS-IV ODD	67	2.45	1.06	.94	58	2.42	0.82	.92	0.03
4. CDRS CD	67	0.24	0.28	.71	58	0.19	0.23	.62	0.16
5. NSIC Callous	67	0.35	0.55	.73	57	0.21	0.37	.67	0.21
6. IRS Impair	67	2.77	2.22	.96	58	2.34	2.01	.94	0.20
7. DBD ADHD	67	1.39	0.97	.98	58	1.12	0.83	.98	0.29
	Pretreatment				Posttreatment				SMD
Teacher Ratings	*N*	*M*	*SD*	α	*N*	*M*	*SD*	α	effect size
8. DBD CP	62	0.69	0.77	.89	46	0.29	0.40	.93	0.76
9. APSD CU	62	56.35	11.29	.71	46	52.26	11.04	.86	0.35
10. ADS-IV ODD	62	2.29	1.12	.95	46	1.88	1.17	.94	0.39
11. CDRS CD	62	0.26	0.41	.82	45	0.12	0.31	.86	0.44
12. NSIC Callous	62	0.62	0.95	.92	45	0.27	0.51	.71	0.53
13. IRS Impair	62	3.00	2.20	.95	46	2.36	2.14	.95	0.30
14. DBD ADHD	62	1.41	0.98	.97	46	1.00	0.84	.97	0.44

Note. SMD effect size = standardized mean difference effect size computed as (pretreatment mean – posttreatment mean)/pooled standard deviation. DBD ADHD, DBD CP = average of ADHD items and average of ODD/CD items on the Disruptive Behavior Disorder Rating Scale; APSD CU = *t* score from the Callous–Unemotional scale of the Antisocial Process Screening Device; ADS-IV ODD = average rating of ODD symptoms on the Assessment of Disruptive Symptoms—DSM-IV version; CDRS CD = average rating of CD symptoms on the Conduct Disorder Rating Scale; NSIC Callous = average rating on the Callous Behavior scale on the Nova Scotia IOWA Conners; IRS Impair = average rating on the Impairment Rating Scale.

that were used in this study are evaluated for frequency over the past 6 months using Likert scales that range from 0 (*never*) to 4 (*daily*). Items were averaged to compute a CD severity score.

Nova Scotia IOWA Connors (NSIC). The NSIC (Waschbusch et al., 2004) consists of 35 items that measure externalizing behavior and peer functioning using Likert scales ranging from 0 (*not at all*) to 3 (*very much*). Of relevance to this study is the Callous Behavior subscale, which consists of the following three items: (a) seems to enjoy being mean, (b) is cold or uncaring, and (c) lacks remorse for misbehavior. These three items were averaged to compute a Callous Behavior score. Previous research on this scale (Waschbusch, Graziano, Willoughby, & Pelham, 2015) has supported its interrater reliability (significant parent–teacher correlation of .40) and validity (significant correlations with the APSD CU scale of .60 in a clinic sample and .50 in a school sample).

Impairment Rating Scale (IRS). The IRS includes seven (parent) or six (teacher) items measuring problem severity and need for treatment in functional domains: relationships with peers, siblings (parent version only), parents (parent version only), and teachers (teacher version only); academic progress (teacher version only); school and classroom behavior; family functioning (parent version only); self-esteem; and overall adjustment (Fabiano et al., 2006). Items were rated using visual-analogue scales anchored by 0 (*no problem; definitely does not need treatment or special services*) and 6 (*extreme problem; definitely does need treatment or special services*) and were averaged into an overall impairment score.

ANALYTIC PLAN

Analyses were intent to treat using all available data. Treatment predictors/moderators in all models were CP measured as the average pretreatment (i.e., baseline) rating of ODD and CD items on the DBDRS, CU measured as pretreatment (baseline) CU scale *t* score on the APSD, informant (parent vs. teacher), treatment (pre vs. post), and their interactions. Medication status during treatment (no vs. yes) and ADHD measured as the average pretreatment (i.e., baseline) rating of ADHD symptoms on the DBDRS were included as covariates. Treatment outcomes were ODD measured as the average rating of ODD items on the ADS-IV, CD measured as the average rating of CD items on the CDRS, callous behavior measured as the average

rating of callousness on the NSIC, and impairment measured as the average rating on the IRS. Thus, separate measures were used for outcomes versus predictors/moderators. In all models, CU and CP were included as continuous measures (centered on sample means) for three reasons: (a) they appear to be continuously distributed in both clinical and general populations (Herpers et al., 2017; Hudziak, Achenbach, Althoff, & Pine, 2007), (b) the observed distributions in this sample approached normality, and (c) continuous predictors are typically more statistically powerful than categorical predictors (Altman & Royston, 2006). Treatment was also a continuous measure, operationalized as the number of weeks before (pretreatment) and after (posttreatment) the first day of treatment, to account for variability across participants regarding the timing of assessments. Significant interactions were probed by computing simple slopes tests at values 1 standard deviation above and below the sample mean and (for interactions involving treatment) at time points 8 weeks before the start of treatment and 8 weeks after the end of treatment. Analyses used full information maximum likelihood procedures to account for missing data (Enders, 2013).

Treatment response was examined using mixed models as computed using Proc Mixed in SAS 9.4, with random intercepts and an unstructured covariance matrix. Kenward–Roger adjustments to degrees of freedom were used to accommodate the small sample size. Treatment outcomes were evaluated by first determining rates of normalization at the end of treatment following methods described by Jacobson and Truax (1991), with normalization defined as having a posttreatment score that was closer to the pretreatment average of control children than to the pretreatment average of children with CP. The cut-score to operationalize normalization in this manner was computed as described by Jacobson and Truax: $(M_{control} \times SD_{cp}) + (M_{cp} \times SD_{controll}) / (SD_{control} + SD_{cp})$, where CP refers to participants with conduct problems and control refers to typically developing participants. Normalization was computed for the full sample of children and also for the subsample of children who had CP (ODD or CD) and normalization rates were compared across parent and teachers by computing odds ratios (see Table 2). Predictors/moderators of normalization (no vs. yes) were examined using logistic regressions computed using Generalized Linear Mixed Models in SPSS version 23.

Table 2
Normalization Rates for the Full Sample and for the Subsample With Pretreatment Conduct Problems Presented Separately by Informant

	Parent ratings				Teacher ratings				Odds ratio
	Not normal		Normal		Not normal		Normal		
Full sample	*n*	%	*n*	%	*n*	%	*n*	%	
DBD CP	36	62.1	22	37.9	19	41.3	27	58.7	2.33*
APSD CU	27	46.6	31	53.4	18	39.1	28	60.9	1.35
ADS ODD	31	53.4	27	46.6	21	45.7	25	54.3	1.37
CDRS CD	21	36.2	37	63.8	10	22.2	35	77.8	1.99
NSIC Callous	20	35.1	37	64.9	14	31.1	31	68.9	1.20
IRS Impairment	29	50.0	29	50.0	21	45.7	25	54.3	1.19
	Parent ratings				Teacher ratings				Odds ratio
	Not normal		Normal		Not normal		Normal		
CP subsample	*n*	%	*n*	%	*n*	%	*n*	%	
DBD CP	30	88.2	4	11.8	18	66.7	9	33.3	3.75*
APSD CU	20	58.8	14	41.2	15	55.6	12	44.4	1.14
ADS ODD	28	82.4	6	17.6	20	74.1	7	25.9	1.63
CDRS CD	18	52.9	16	47.1	7	26.9	19	73.1	3.05*
NSIC Callous	17	50.0	17	50.0	12	46.2	14	53.8	1.17
IRS Impairment	25	73.5	9	26.5	17	63.0	10	37.0	1.63

Note. Values in the table are the number and percentage of participants who were or were not within the normal range of functioning at the end of treatment. Odds ratio = odds of normalization according to teacher ratings versus parent ratings. DBD CP = average rating of ODD/CD items on the Disruptive Behavior Disorder Rating Scale; APSD CU = *t* score from the Callous–Unemotional scale of the Antisocial Process Screening Device; ADS-IV ODD = average rating of ODD symptoms on the Assessment of Disruptive Symptoms—DSM-IV version; CDRS CD = average rating of CD symptoms on the Conduct Disorder Rating Scale; NSIC Callous = average rating on the Callous Behavior scale on the Nova Scotia IOWA Conners; IRS Impairment = average rating on the Impairment Rating Scale.
* = different from zero at $p < .05$.

Results

TREATMENT RESPONSE

Oppositional Defiant Disorder

There were significant main effects of CP, $F(1, 67.6) = 31.26$, $p < .0001$; CU, $F(1, 69.1) = 4.35$, $p = .0407$; and informant, $F(1, 166) = 10.35$, $p = .0016$, and a marginal effect of treatment, $F(1, 188) = 3.89$, $p = .0501$. The main effect of treatment showed ODD scores decreased with treatment (b = –.01) and the informant main effect showed that ODD scores, averaged across pre- and posttreatment, were higher for parent than teacher (parent $M = 2.46$, $SE = 0.07$; teacher $M = 2.12$, $SE = 0.08$). The main effects of CP and CU were qualified by a significant CP × CU interaction, $F(1, 66.8) = 7.01$, $p = .0101$. As shown in Figure 1, in participants with lower CP before treatment, higher CU was significantly associated with higher ODD.

Conduct Disorder

There were significant main effects of CP, $F(1, 66.1) = 39.74$, $p < .0001$, and informant, $F(1, 165) = 4.08$, $p = .0449$, but these were qualified by significant CP × CU × Treatment, $F(1, 199) = 5.96$, $p = .0155$, and CP × CU × Informant, $F(1, 161) = 6.03$, $p = .0152$, interactions. For ease of interpretation, the two interactions were graphed and followed up together (see Figure 2). Follow-up tests showed that pretreatment CU significantly moderated treatment response for both parent- and teacher-rated CD. As shown in Figure 2, in participants with higher pretreatment CP, treatment response was larger for participants with higher CU (i.e., parents and teachers reported significant reductions in CD for participants with higher CP and CU).

Callous Behavior

There were significant main effects of CP, $F(1, 53.8) = 16.76$, $p = .0001$, and CU, $F(1, 56.4) = 15.73$, $p =$

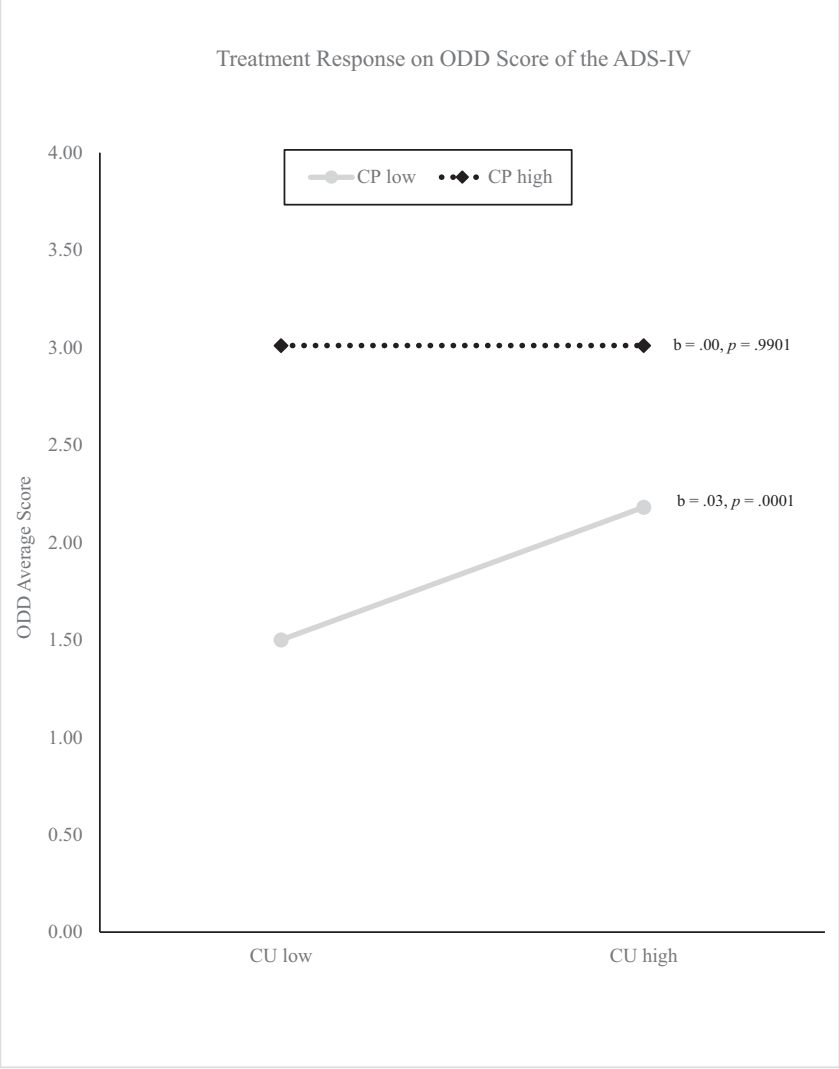

FIGURE 1 Oppositional defiant disorder (ODD) average scores (range = 0–4; higher scores indicate more serious ODD) as a function of pretreatment CP and pretreatment CU.

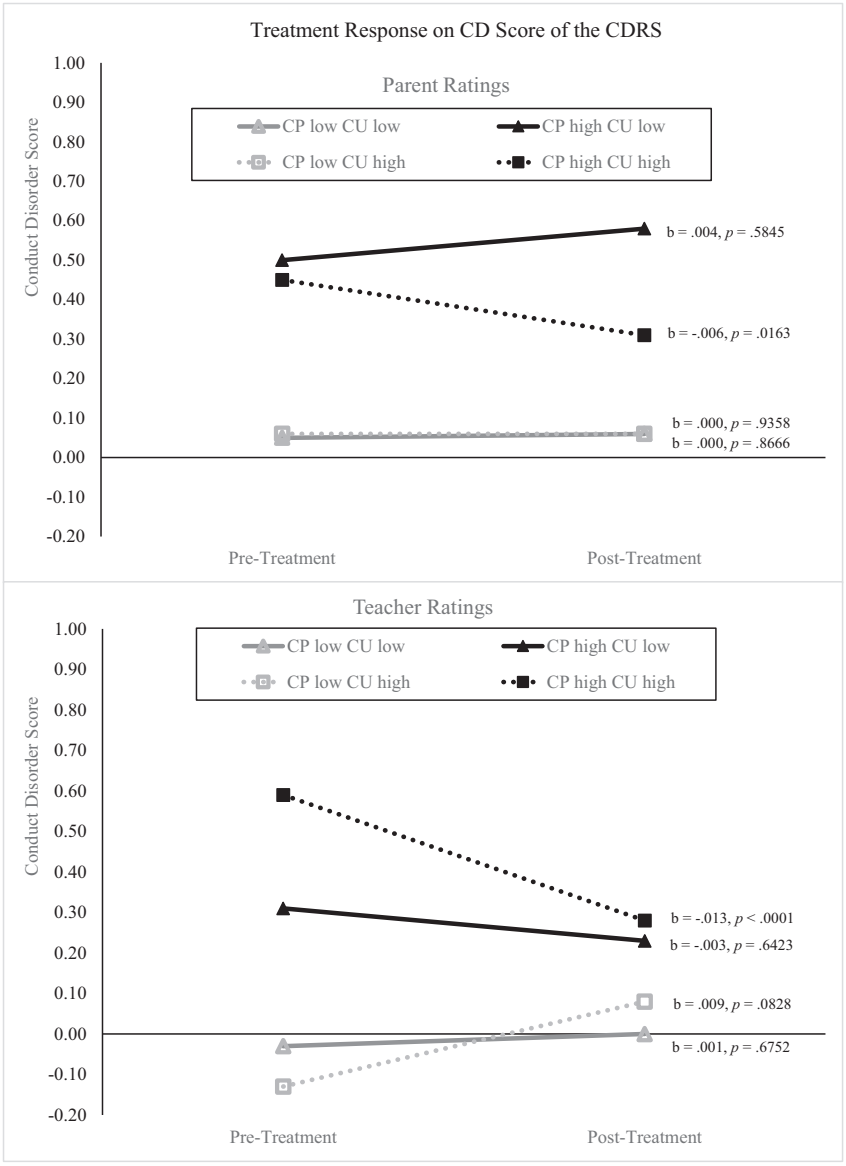

FIGURE 2 Conduct disorder average scores (range = 0–4; higher scores indicate more serious CD) as a function of informant, treatment, pretreatment CP, and pretreatment CU. Some values are less than zero due to adjustments from covariates (medication, ADHD).

.0002, as well as significant interactions among CP × CU, $F(1, 54.4) = 5.51$, $p = .0226$, CU × Informant, $F(1, 149) = 7.44$, $p = .0072$, CP × CU × Informant, $F(1, 150) = 8.12$, $p = .0050$, and CU × Treatment, $F(1, 176) = 5.00$, $p = .0265$, but these were qualified by a significant CP × CU × Informant × Treatment interaction, $F(1, 156) = 6.15$, $p = .0142$. Follow-up tests showed that pretreatment CU moderated treatment response for teacher ratings but not parent ratings. As shown in Figure 3, in participants with higher pretreatment CP, treatment response was greater for participants with higher CU (i.e., teacher-rated Callous Behavior scores were significantly reduced for participants with higher CP and higher CU).

Impairment

There were significant main effects of CP, $F(1, 58) = 11.70$, $p = .0012$, and treatment, $F(1, 182) = 8.60$, $p = .0038$, but these were qualified by a significant CP × Treatment interaction, $F(1, 177) = 5.24$, $p = .0233$. Follow-up tests of the interaction showed that treatment decreased impairment for participants who had higher pretreatment CP (pretreatment $M = 3.77$, $SE = 0.30$; posttreatment $M = 2.78$, $SE = 0.31$; $b_{treatment} = -.04$, $p = .0009$) but there was no change in impairment for participants with lower pretreatment CP (pretreatment $M = 2.12$, $SE = 0.23$; posttreatment $M = 2.08$, $SE = 0.24$; $b_{treatment} = -.002$, $p = .8537$).

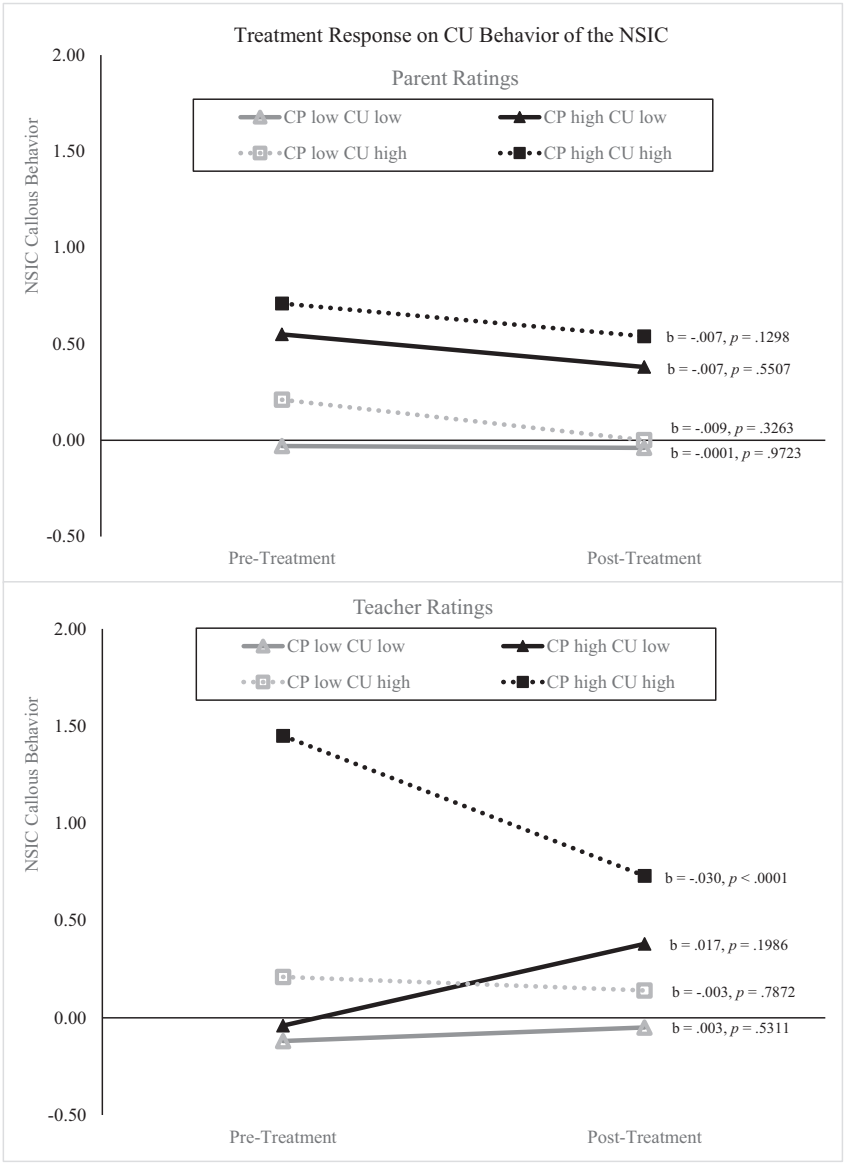

FIGURE 3 Callous Behavior scores (range = 0–3; higher scores indicate more serious callous behavior) as a function of informant, treatment, pretreatment CP, and pretreatment CU. Some values are less than zero due to adjustments from covariates (medication, ADHD).

TREATMENT OUTCOME

Oppositional Defiant Disorder

Odds ratios showed normalization rates did not differ across parent and teacher for either the full sample or the CP subsample (see Table 2). Logistic regressions showed a main effect of CP, $F(1, 94) = 4.09$, $p = .046$, which showed higher pretreatment CP was associated with a lower likelihood of posttreatment normalization (b = –1.79).

Conduct Disorder

Odds ratios showed normalization rates were 1.99 times higher in the full sample and 3.05 times higher in the CP subsample for teacher- versus parent-rated CD outcomes (see Table 2). Logistic regressions showed significant main effects of ADHD, $F(1, 93) = 7.50$, $p = .007$, CP, $F(1, 93) = 10.12$, $p = .002$], and informant, $F(1, 93) = 8.44$, $p = .005$, as well as the following interactions: CP × CU, $F(1, 93) = 5.14$, $p = .026$, CP × Informant, $F(1, 93) = 5.92$, $p = .017$, and CU × Informant, $F(1, 93) = 7.59$, $p = .007$. For ease of interpretation, the interactions were graphed and followed up together (see Figure 4). Follow-up tests showed that pretreatment CU moderated the association between pretreatment CP for teacher ratings of CD but not parent ratings

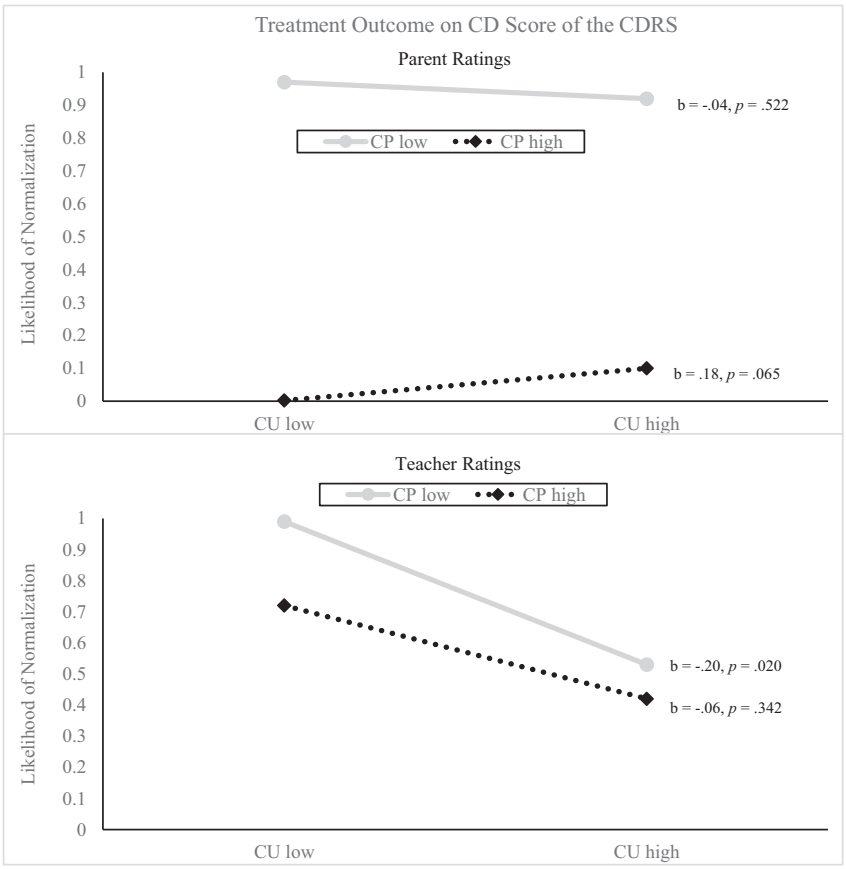

FIGURE 4 Likelihood of normalization on conduct disorder scores as a function of informant, pretreatment CP, and pretreatment CU.

of CD. As shown in Figure 4, in patients with lower CP before treatment, treatment outcomes were worse for higher levels of CU (i.e., teacher ratings of CD showed a lower likelihood of normalization for participants with lower CP and higher CU before treatment).

Callous Behavior

Odds ratios showed no differences in rates of normalization for parent versus teacher ratings (see Table 2). Logistic regressions showed a significant main effect of CU, $F(1, 92) = 6.21$, $p = .014$, and a marginal main effect of CP, $F(1, 92) = 2.91$, $p = .091$. These results showed that higher pretreatment CU (b = –.14) and higher pretreatment CP (b = –1.29) were each associated with a lower likelihood of normalization.

Impairment

Odds ratios showed no differences in rates of normalization for parent versus teacher ratings (see Table 2). Logistic regressions showed a significant main effect of ADHD, $F(1, 94) = 5.84$, $p = .018$, which showed that higher ADHD was associated with a lower likelihood of normalization (b = –1.01).

Discussion

The overarching purpose of this study was to examine whether the association between CP and treatment was moderated by CU traits. Treatment consisted of BT delivered to children in a summer camp setting and to parents in weekly parent training. Both treatment response and treatment outcome were examined, with the former operationalized as change between pre- and posttreatment and the latter operationalized as the likelihood of normalization after treatment. Informant differences in treatment evaluations were also examined by directly comparing parent and teacher ratings. It was hypothesized that treatment response and treatment outcome would each be associated with pretreatment CP, but the association would be moderated by pretreatment CU. Specifically, high CP and high CU at pretreatment were expected to be associated with a better treatment *response* but worse treatment *outcome* relative to high CP and low CU at pretreatment. Treatment effects were also expected to be larger for parent than teacher ratings.

The first set of analyses examined treatment response and showed that CU moderated the

association between CP and treatment response on two of four measures. Specifically, treatment significantly reduced CD scores for participants with higher CP and higher CU before treatment, and this was found for both parent and teacher ratings of CD (see Figure 2). Likewise, Callous Behavior scores were significantly reduced for participants with higher CP and higher CU before treatment, although this was true only for teacher ratings of CU (see Figure 3). Thus, as hypothesized, the combination of high CP and high CU was associated with a larger treatment response on these measures. In contrast, CU did not moderate treatment response on ODD scores or on impairment scores. Instead, there was a marginal (p = .0501) reduction in ODD scores for the entire sample, as well as reduced impairment for participants with higher CP before treatment. These latter findings did not support our hypotheses. Overall, there was mixed support for our hypothesis that CU would moderate the association between CP and treatment response.

The second set of analyses examined treatment outcome, defined as posttreatment normalization. Among participants with CP, rates of normalization after treatment averaged 35% for parent-rated outcomes and 47% for teacher-rated outcomes (see Table 2). These rates are in line with the 40% normalization rate reported in a review of parent-training studies (Sheldrick, Kendall, & Heimberg, 2001) and with the 42% normalization rate reported in one STP study (Pelham et al., 2000). More directly relevant to this study, the treatment outcomes analyses also showed that both CU and informant moderated the association between CP and treatment outcome on one of four measures. As shown in the top of Figure 4, CU did not significantly moderate the association between CP and parent ratings of CD—the likelihood of normalization on parent-rated CD was very high for participants with low CP before treatment but very low for participants with high CP before treatment. However, as shown in the bottom of Figure 4, CU did moderate the association between pretreatment CP and teacher-rated CD. Specifically, only participants with lower CU and lower CP before treatment had very high rates of normalization after treatment; participants with high CP or high CU before treatment had markedly lower rates of posttreatment normalization. Other than this one measure, CU did not moderate treatment outcomes. Instead, higher pretreatment CP (regardless of the level of pretreatment CU) was associated with worse outcomes on ODD and callous behavior, and higher pretreatment CU (regardless of pretreatment CP) was associated with worse outcomes on callous behavior. Overall, these results showed that normalization rates were generally lower for children with higher CP before treatment, and this pattern did not differ as a function of CU (with one exception—teacher ratings of CD). Taken as a whole, these findings are consistent with previous studies (Lindhiem et al., 2012; Reyno & McGrath, 2006), suggesting that those with more severe problems before treatment are less likely to be normalized after treatment, but our results add to these by suggesting that rates of normalization generally do not differ for CP children with and without CU traits.

With regard to informant differences in evaluating treatment effects, two noteworthy findings emerged. First, treatment effects were larger as measured by teacher ratings than parent ratings. Specifically, the magnitude of change between pre- and posttreatment (treatment response) was, on average, twice as large according to teacher ratings than parent ratings (see effect sizes in Table 1) and the odds of normalization from treatment were also higher for teacher ratings at posttreatment (see Table 2). These findings run counter to what is typically reported on treatment studies, which tend to show larger treatment effects on parent ratings than on teacher ratings (Weisz et al., 2017). The reasons for this difference are not clear, but one speculation is that the STP includes 3 hours per day of classroom time focused specifically on improving classroom behavior and performance. This is a more direct and intense intervention than is available in most other treatment programs. Second, teacher ratings made important contributions to understanding the effects of CU on the association between CP and treatment. As shown in Figures 2–4, effects of CU on treatment response and treatment outcome sometimes differed between parent and teacher ratings, and failure to include teacher ratings may have missed detecting these treatment effects. This is an important point because the majority of CU treatment studies have not included teacher-rated outcomes when examining treatment effects.

There are a few important limitations to note when interpreting the results. First, the sample size was modest, although it is similar to or larger than many previous CU treatment studies (Haas et al., 2011; McDonald, Dodson, Rosenfield, & Jouriles, 2011; Waschbusch et al., 2007). The primary weakness of a small sample size is limited statistical power to detect treatment effects, but the within-subjects design and continuous variables helps to mitigate these concerns. Second, though the sample was reflective of the community from which participants were recruited, the findings may not generalize to other communities that have more racially and ethnically diverse populations. Third,

different teachers completed ratings at pre- versus posttreatment. We do not believe this is a serious shortcoming because the posttreatment teachers completed ratings well after the start of the school year (an average of 11 weeks after the end of treatment and 10 weeks after the start of school). This is more than sufficient time for the teacher to get to know students very well, considering that (a) some behavior ratings specify 4 weeks as an adequate time frame (e.g., the Devereaux Scale of Mental Disorders; Naglieri, LeBuffe, & Pfeiffer, 1994), and (b) disruptive students often differentiate themselves from typically developing students within hours or even minutes of being introduced to others (Erhardt & Hinshaw, 1994; Pelham & Bender, 1982).

Despite these limitations, the current study meaningfully extends research on the treatment of children with CP and CU, such that these children who receive intensive BT are likely to improve yet will likely continue needing support and treatment. These findings reinforce the notion that children with especially serious mental health problems, such as children with CP and CU traits, are likely to need intensive and specialized treatment; traditional outpatient treatment is unlikely to be sufficient for such seriously impaired children (Comer & Barlow, 2014).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.beth.2018.03.003.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037/0033-2909.101.2.213

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, DC: Author.

Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, *27*, 180–189. https://doi.org/10.1207/s15374424jccp2702_5

Burns, G. L., Walsh, J. A., Servera, M., Lorenzo-Seva, U., Cardo, E., & Rodriguez-Fornells, A. (2013). Construct validity of ADHD/ODD rating scales: Recommendations for the evaluation of forthcoming DSM-V ADHD/ODD scales. *Journal of Abnormal Child Psychology*, *41*, 15–26. https://doi.org/10.1007/s10802-012-9660-5

Comer, J. S., & Barlow, D. H. (2014). The occasional case against broad dissemination and implementation: Retaining a role for specialty care in the delivery of psychological treatments. *American Psychologist*, *69*, 1–18. https://doi.org/10.1037/a0033582

Comer, J. S., Chow, C., Chan, P. T., Cooper-Vince, C., & Wilson, L. A. (2013). Psychosocial treatment efficacy for disruptive behavior problems in very young children: A meta-analytic examination. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*, 26–36. https://doi.org/10.1016/j.jaac.2012.10.001

De Los Reyes, A. (2011). Introduction to the special section. More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *40*, 1–9. https://doi.org/10.1080/15374416.2011.533405

Enders, C. K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives*, *7*, 27–31. https://doi.org/10.1111/cdep.12008

Erhardt, D., & Hinshaw, S. P. (1994). Initial sociometric impressions of attention-deficit hyperactivity disorder and comparison boys: Predictions from social behaviors and from nonbehavioral variables. *Journal of Consulting and Clinical Psychology*, *62*, 833–842. https://doi.org/10.1037/0022-006X.62.4.833

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*, *37*, 215–237. https://doi.org/10.1080/15374410701820117

Fabiano, G. A., Pelham, W. E., Jr., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., . . . MacLean-Burrows, L. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology*, *35*, 369–385. https://doi.org/10.1207/s15374424jccp3503_3

Frick, P. J., & Hare, R. D. (2001). *Antisocial Process Screening Device: Technical manual*. Toronto, Canada: Multi-Health Systems.

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous–unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, *140*, 1–57. https://doi.org/10.1037/a0033076

Haas, S. M., Waschbusch, D. A., Pelham, W. E., Jr., King, S., Andrade, B. F., & Carrey, N. J. (2011). Treatment response in CP/ADHD children with callous/unemotional traits. *Journal of Abnormal Child Psychology*, *39*, 541–552. https://doi.org/10.1007/s10802-010-9480-4

Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous–unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical Child and Family Psychology Review*, *17*, 248–267. https://doi.org/10.1007/s10567-014-0167-1

Herpers, P. C. M., Klip, H., Rommelse, N. N. J., Taylor, M. J., Greven, C. U., & Buitelaar, J. K. (2017). Taxometric analyses and predictive accuracy of callous–unemotional traits regarding quality of life and behavior problems in non-conduct disorder diagnoses. *Psychiatry Research*, *253*, 351–359. https://doi.org/10.1016/j.psychres.2017.04.004

Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, *16*(Suppl. 1), S16–S23. https://doi.org/10.1002/mpr.217

Hyde, L. W., Shaw, D. S., Gardner, F., & Cheong, J. (2013). Dimensions of callousness in early childhood: Links to problem behavior and family intervention effectiveness. *Development and Psychopathology*, *25*, 347–363. https://doi.org/10.1017/S0954579412001101

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Kolko, D. J. (1995). Multimodal partial/day treatment of child antisocial behavior: Service description and multilevel program evaluation. *Continuum: Developments in Ambulatory Mental Health Care*, *2*, 3–24.

Kolko, D. J., Bukstein, O. G., & Barron, J. (1999). Methylphenidate and behavior modification in children with ADHD and comorbid ODD or CD: Main and incremental effects across settings. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 578–586. https://doi.org/10.1097/00004583-199905000-00020

Kolko, D. J., Dorn, L. D., Bukstein, O. G., Pardini, D. A., Holden, E. A., & Hart, J. (2009). Community vs. clinic-based modular treatment of children with early-onset ODD or CD: A clinical trial with 3-year follow-up. *Journal of Abnormal Child Psychology*, *37*, 591–609. https://doi.org/10.1007/s10802-009-9303-7

Kolko, D. J., & Pardini, D. A. (2010). ODD dimensions, ADHD, and callous–unemotional traits as predictors of treatment response in children with disruptive behavior disorders. *Journal of Abnormal Psychology*, *119*, 713–725. https://doi.org/10.1037/a0020910

Lindhiem, O., Kolko, D. J., & Cheng, Y. (2012). Predicting psychotherapy benefit: A probabilistic and individualized approach. *Behavior Therapy*, *43*, 381–392. https://doi.org/10.1016/j.beth.2011.08.004

Loeber, R., Green, S. M., Lahey, B. B., Christ, M. A. G., & Frick, P. J. (1992). Developmental sequences in the age of onset of disruptive child behaviors. *Journal of Child and Family Studies*, *1*, 21–41. https://doi.org/10.1007/BF01321340

McDonald, R., Dodson, M. C., Rosenfield, D., & Jouriles, E. N. (2011). Effects of a parenting intervention on features of psychopathy in children. *Journal of Abnormal Child Psychology*, *39*, 1013–1023. https://doi.org/10.1007/s10802-011-9512-8

McMahon, R. J., & Frick, P. J. (2005). Evidence-based assessment of conduct problems in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 477–505. https://doi.org/10.1207/s15374424jccp3403_6

Miller, N. V., Haas, S. M., Waschbusch, D. A., Willoughby, M. T., Helseth, S. A., Crum, K. I., . . . Pelham, W. E., Jr. (2014). Behavior therapy and callous–unemotional traits: Effects of a pilot study examining modified behavioral contingencies on child behavior. *Behavior Therapy*, *45*, 606–618. https://doi.org/10.1016/j.beth.2013.10.006

Naglieri, J. A., LeBuffe, P. A., & Pfeiffer, S. I. (1994). *The Devereaux Scales of Mental Disorders*. San Antonio, TX: Harcourt Assessment.

Ollendick, T. H., Greene, R. W., Austin, K. E., Fraire, M. G., Halldorsdottir, T., Allen, K. B., . . . Wolff, J. C. (2015). Parent management training and collaborative and proactive solutions: A randomized control trial for oppositional youth. *Journal of Clinical Child and Adolescent Psychology*, *45*, 591–604. https://doi.org/10.1080/15374416.2015.1004681

Pelham, W. E., Jr., & Bender, M. E. (1982). Peer relationships in hyperactive children: Description and treatment. In B. G. Cook, M. Tankersley, & Landrum, T. J. (Eds.), *Advances in learning and behavioral disorders* (Vol. 1, pp. 365–436). Greenwich, CT: JAI Press.

Pelham, W. E., Jr., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 449–476. https://doi.org/10.1207/s15374424jccp3403_5

Pelham, W. E., Jr., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 210–218. https://doi.org/10.1097/00004583-199203000-00006

Pelham, W. E., Jr., Gnagy, E. M., Greiner, A. R., Fabiano, G. A., Waschbusch, D. A., & Burrows-MacLean, L. (2010). Summer treatment programs for attention-deficit/hyperactivity disorder. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidenced-based psychotherapies for children and adolescents* (2nd ed., pp. 277–292). New York, NY: Guilford Press.

Pelham, W. E., Jr., Gnagy, E. M., Greiner, A. R., Hoza, B., Hinshaw, S. P., Swanson, J. M., . . . McBurnett, K. (2000). Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer treatment program. *Journal of Abnormal Child Psychology*, *28*, 507–525. https://doi.org/10.1023/A:1005127030251

Pelham, W. E., Jr., Greiner, A. R., Gnagy, E. M., Hoza, B., Martin, L., Sams, S., & Wilson, T. (1996). Intensive treatment for ADHD: A model summer treatment program. In M. C. Roberts (Ed.), *Model programs in child and family mental health* (pp. 193–213). New York, NY: Psychology Press.

Pelham, W. E., Jr., & Hoza, B. (1996). Intensive treatment: A summer treatment program for children with ADHD. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychological treatment for child and adolescent disorders: Empirically based strategies for clinical practice* (pp. 311–340). Washington, DC: American Psychological Association.

Reid, M. J., Webster-Stratton, C., & Hammond, M. (2003). Follow-up of children who received the Incredible Years intervention for oppositional-defiant disorder: Maintenance and prediction of 2-year outcome. *Behavior Therapy*, *34*, 471–491. https://doi.org/10.1016/S0005-7894(03)80031-X

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems: A meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*, 99–111. https://doi.org/10.1111/j.1469-7610.2005.01544.x

Sheldrick, R. C., Kendall, P. C., & Heimberg, R. G. (2001). The clinical significance of treatments: A comparison of three treatments for conduct disordered children. *Clinical Psychology: Science and Practice*, *8*, 418–430. https://doi.org/10.1093/clipsy.8.4.418

Stattin, H., Enebrink, P., Özdemir, M., & Giannotta, F. (2015). A national evaluation of parenting programs in Sweden: The short-term effects using an RCT effectiveness design. *Journal of Consulting and Clinical Psychology*, *83*, 1069–1084. https://doi.org/10.1037/a0039328

Waller, R., Gardner, F., & Hyde, L. W. (2013). What are the associations between parenting, callous–unemotional traits, and antisocial behavior in youth? A systematic review of evidence. *Clinical Psychology Review*, *33*, 593–608. https://doi.org/10.1016/j.cpr.2013.03.001

Waschbusch, D. A. (2002). A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems. *Psychological Bulletin*, *128*, 118–150. https://doi.org/10.1037/0033-2909.128.1.118

Waschbusch, D. A., Bernstein, M. D., Robb Mazzant, J., Willoughby, M. T., Haas, S. M., Coles, E. K., & Pelham, W. E., Jr. (2016). A case study examining fixed versus randomized criteria for treating a child with conduct problems

and callous–unemotional traits. *Evidenced-Based Practice in Child and Adolescent Mental Health*, *1*, 73–85. https://doi.org/10.1080/23794925.2016.1227946

Waschbusch, D. A., Carrey, N. J., Willoughby, M. T., King, S., & Andrade, B. F. (2007). Effects of methylphenidate and behavior modification on the social and academic behavior of children with disruptive behavior disorders: The moderating role of callous/unemotional traits. *Journal of Clinical Child and Adolescent Psychology*, *36*, 629–644. https://doi.org/10.1080/15374410701662766

Waschbusch, D. A., & Elgar, F. J. (2007). Development and validation of the Conduct Disorder Rating Scale. *Assessment*, *14*, 65–74. https://doi.org/10.1177/1073191106289908

Waschbusch, D. A., Graziano, P. A., Willoughby, M. T., & Pelham, W. E., Jr. (2015). Classroom rule violations in elementary school students with callous–unemotional traits. *Journal of Emotional and Behavioral Disorders*, *23*, 180–192. https://doi.org/10.1177/1063426614552903

Waschbusch, D. A., Porter, S., Carrey, N. J., Kazmi, S. O., Roach, K. A., & D'Amico, D. A. (2004). Investigation of the heterogeneity of disruptive behaviour in elementary-age children. *Canadian Journal of Behavioural Science*, *36*, 97–112. https://doi.org/10.1037/h0087221

Waschbusch, D. A., & Sparks, S. J. (2003). Rating scale assessment of attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD): Is there a normal distribution and does it matter? *Journal of Psychoeducational Assessment*, *21*, 261–281. https://doi.org/10.1177/073428290302100303

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children - Third Edition.* New York: Psychological Coporation.

Weisz, J. R., Sofie, K., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., . . . Forwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, *72*, 79–117. https://doi.org/10.1037/a0040360

Wilkinson, S., Waller, R., & Viding, E. (2016). Practitioner review. Involving young people with callous unemotional traits in treatment: Does it work? A systematic review. *Journal of Child Psychology and Psychiatry*, *57*, 552–565. https://doi.org/10.1111/jcpp.12494

Received: July 30, 2017
Accepted: March 5, 2018
Available online: 8 March 2018