



## Behavior Therapy and Callous-Unemotional Traits: Effects of a Pilot Study Examining Modified Behavioral Contingencies on Child Behavior

Natalie V. Miller

University of British Columbia

Sarah M. Haas

Center for Children and Families, University of Buffalo–SUNY

Daniel A. Waschbusch

Center for Children and Families, Florida International University  
Penn State Hershey Medical Center

Michael T. Willoughby

Frank Porter Graham Child Development Institute  
The University of North Carolina at Chapel Hill

Sarah A. Helseth

Kathleen I. Crum

Erika K. Coles

William E. Pelham, Jr.

Center for Children and Families, Florida International University

The conduct problems of children with callous-unemotional (CU) traits (i.e., lack of empathy, lack of guilt/lack of caring behaviors) are particularly resistant to current behavioral interventions, and it is possible that differential sensitivities to punishment and reward may underlie this resistance. Children with conduct problems and CU (CPCU) are less responsive to

behavioral punishment techniques (e.g., time-out), whereas reward techniques (e.g., earning points for prizes or activities) are effective for reducing conduct problems. This study examined the efficacy of modified behavioral interventions, which de-emphasized punishment (Condition B) and emphasized reward techniques (Condition C), compared with a standard behavioral intervention (Condition A). Interventions were delivered through a summer treatment program over 7 weeks with an A-B-A-C-A-BC-A design to a group of 11 children (7–11 years; 91% male). All children were diagnosed with either oppositional defiant disorder or conduct disorder, in addition to attention-deficit hyperactivity disorder. Results revealed the best treatment response occurred during the low-punishment condition, with rates of negative behavior (e.g., aggression, teasing, stealing) increasing over the 7 weeks. However, there was substantial individual variability in treatment response, and several children demonstrated

This study was supported by the National Institute of Mental Health (7R34MH085796) awarded to Daniel Waschbusch. The authors also wish to thank the staff at the Center for Children and Families, and the families that participated in the study.

Address correspondence to Natalie V. Miller, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia, Canada V6T 1Z4; e-mail: [nvmiller@psych.ubc.ca](mailto:nvmiller@psych.ubc.ca).

0005-7894/xx/xxx-xxx/\$1.00/0

© 2013 Association for Behavioral and Cognitive Therapies. Published by Elsevier Ltd. All rights reserved.

improvement during the modified intervention conditions. Future research is necessary to disentangle treatment effects from order effects, and implications of group treatment of CPCU children (i.e., deviancy training) are discussed.

---

*Keywords:* callous-unemotional traits; conduct problems; attention-deficit/hyperactivity disorder; summer treatment program

APPROXIMATELY 10% OF SCHOOL-AGE children have significant conduct problems, including aggression, disruptive behavior, and verbal conflict (Nock, Kazdin, Hiripi, & Kessler, 2006, 2007). In the absence of intervention, these problems tend to persevere and are linked to poorer outcomes in adolescence (e.g., substance abuse) and adulthood (e.g., incarceration; Moffitt, Caspi, Dickson, Silva, & Stanton, 1996; Robins & Price, 1991). Behavioral interventions have received strong empirical support for improving the current conduct problems (Wilson, Lipsey, & Derzon, 2003) and long-term trajectories of these children (Conduct Problems Prevention Research Group, 2011). However, nearly one third of treated children do not benefit from these interventions (Masi et al., 2011; Webster-Stratton & Hammond, 1997). In an effort to better understand treatment response, recent research has focused on identifying characteristics that may account for this difference.

One constellation of child characteristics that may be relevant in this regard are callous-unemotional (CU) traits, which refers to an uncaring, unemotional, and callous affective style and a conning, manipulative, and self-serving interpersonal style (Frick et al., 2003; Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007). Research published over the last two decades demonstrates that CU traits explain significant variance within the population of children with conduct problems (CP). Children with CP and CU traits (CPCU), as compared with children with conduct problems without CU traits (CP-only), are more likely to exhibit severe and violent antisocial behavior (Frick & Ellis, 1999; McMahon, Witkiewitz, & Kotler, 2010). Based on these and many other studies, *with limited prosocial emotions* (i.e., CU traits) has been added as a specifier of conduct disorder in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013).

Importantly, there is some evidence that behavioral treatment for conduct problems—which is considered a “best-practice” approach (Eyberg, Nelson, & Boggs, 2008)—may be insufficiently effective for children with elevated levels of CU traits (Haas et al., 2011; Hawes & Dadds, 2005; Kimonis &

Armstrong, 2012; Masi et al., 2011; Waschbusch, Walsh, Andrade, King, & Carrey, 2007). That is, the conduct problems of children with CU traits are more likely to persist following behavioral treatment compared to children with CP-only, and this result has been reported for both parent- and child-based interventions (e.g., Haas et al., 2011; Hawes & Dadds, 2005). The relation between CU traits and poor treatment response has been found even after controlling for preexisting conduct problems, suggesting that it is CU traits themselves, rather than the associated conduct problem severity, that contribute to poor treatment response (Waschbusch et al., 2007).

If CU traits are associated with decreased response to behavioral interventions, what might account for this finding? One possibility is that this pattern may reflect differences in punishment and/or reward sensitivity. In particular, it has been hypothesized that children with CU traits are underresponsive to punishment and overfocused on reward (Dadds & Salmon, 2003). This hypothesis is supported by results from studies using controlled experimental tasks, which have found that children with CPCU are less likely to change their behavior in response to punishment (typically operationalized as a loss of points), especially when they are first primed with a reward (Blair, Colledge, & Mitchell, 2001; O'Brien & Frick, 1996). Further, functional magnetic resonance imaging research shows these behavioral differences are accompanied by differences in brain activation, suggesting that children with CPCU process punishment and reward decisions in qualitatively distinct ways (Finger et al., 2008).

There also is evidence, albeit indirect evidence, from treatment studies that suggests children with CPCU are underresponsive to punishment, overfocused on reward, or both. Time-out procedures, arguably a punishment technique, may be less effective for children with CPCU than for children with CP-only (Haas et al., 2011; Hawes & Dadds, 2005). Haas and colleagues examined response to intensive behavioral treatment among elementary school-age children with CP and attention-deficit hyperactivity disorder (ADHD) and reported that higher CU traits were associated with more negative behavior during time-out. On the other hand, reward-based strategies may be effective for reducing problematic behavior among children with CPCU (Hawes & Dadds, 2005).

The implication of these findings is clear; standard behavior therapy, in which punishment and reward techniques are used in a balanced fashion to shape behavior, may not be as effective for children with CPCU. If so, implementing behavior therapy in a manner that de-emphasizes punishment and emphasizes reward techniques may be a more

effective approach for children with CPCU. In theory, this latter approach may align more closely with the unique punishment insensitive/reward-focused learning style exhibited by children with CPCU.

The limited available treatment research suggests that punishment techniques used in standard behavioral treatments (i.e., time-out, loss of privileges) are not only ineffective (Hawes & Dadds, 2005) but may actually induce escalations in negative behaviors among children with CU traits (Haas et al., 2011). Therefore, reducing (as much as possible) punishment by ignoring as much inappropriate behavior as possible, while recognizing that not all negative behavior can be ignored, may be an optimal approach for children with CPCU. Further, to the extent that CU traits are associated with goal-directed/reward-driven behavior regardless of the negative consequences, and to the extent that CU traits are associated with being self-serving, then it also may be effective to use reward-based strategies to shape the behaviors of children with CPCU. In other words, strong and frequent rewards may be a tool that can align the child's heightened self-interest with the interests of those around him or her.

The purpose of this study was to use single-case research methods to examine the response of children with CPCU to four behavioral treatment conditions: (a) standard treatment in which punishment and reward were delivered in a balanced manner; (b) low-punishment treatment, in which negative consequences for undesirable/negative behaviors were minimized; (c) high-reward treatment, in which positive consequences for positive or nonnegative behavior were maximized; and (d) combined treatment, in which punishments were minimized and rewards were maximized. We chose a single-case design for its versatility, as we would be implementing several different treatment conditions over a short period of time to a small group of children. We hypothesized that rates of negative behavior (i.e., noncompliance, conduct problems, negative verbalizations) would be significantly lower in the low-punishment, high-reward, and combined conditions compared with standard treatment. Further, we predicted that negative behavior would be lowest in the combined condition, and that the low-punishment and high-reward conditions would perform equally well. As a secondary aim, we were interested in whether children's behavior during time-out would change across conditions, as time-out has been identified in previous research as an area of particular difficulty for behavioral interventions with this population. Parallel to our primary hypotheses, we

expected a similar pattern of treatment response across conditions.

## Method

### PARTICIPANTS

Participants were 11 children (1 female) between the ages of 7 and 11 ( $M = 9.5$ ,  $SD = 1.2$ ). Participants were recruited between April and June 2011 from a larger treatment program for children with ADHD, oppositional defiant disorder (ODD), and conduct disorder (CD). Families were recruited to the larger program using radio and television advertisements, postings in public places (e.g., university websites, billboards), fliers distributed to pediatrician offices and elementary schools, and word of mouth from families enrolled in previous years. Of the 76 participants in the larger treatment program, 35 met eligibility criteria for this research study (i.e., IQ of 80 or above, between 7 and 11 years old, diagnosed with ODD or CD, and had clinically significant levels of CU traits). The remaining children were not in the appropriate age range ( $n = 22$ ), did not meet criteria for ODD or CD ( $n = 15$ ), or did not have clinically significant levels of CU traits ( $n = 4$ ). Families first completed an assessment for the larger treatment program and eligibility to participate in this research study was determined based on the information gathered at the initial assessment. Eligible families were contacted and given information about the nature and purpose of the research program and they were offered the chance to enroll their child in the research program in exchange for receiving the treatment free of charge. This process was conducted on a first-come-first-served basis until 12 participants were enrolled in the study. In total, 15 potentially eligible families were approached with 3 declining participation. Two families declined for pragmatic reasons, and 1 family declined because they were not interested in the project. One child, who was enrolled in the study, was dropped from data analyses due to noncompliance with the research protocol (parents administered psychoactive medication to the child, which was prohibited). Children were predominantly Hispanic (83%). Average household income ranged from \$15,000 to \$59,000 ( $M = \$36,000$ ,  $SD = \$14,433$ ). Eighty-two percent of the children lived with their biological parents (45% married/common law, 36% divorced/separated/single), and 18% lived with relatives or adoptive parents. Intellectual abilities, as assessed using the WISC-IV, ranged from full scale IQ of 88 to 123 ( $M = 104$ ,  $SD = 12$ ).

All children met criteria for conduct problems, defined as ODD or CD, as well as ADHD. Diagnoses were made by Ph.D. clinicians using several sources of information, including parent and teacher ratings on

Table 1  
Descriptive Statistics for Disruptive Behavior Disorders and  
CU Traits Measures

| Measure                       | Parent     | Teacher     |
|-------------------------------|------------|-------------|
| DBDSPI symptom counts         |            |             |
| Hyperactivity/impulsivity     | 5.0 (1.4)  | --          |
| Inattention                   | 6.6 (2.3)  | --          |
| Oppositional defiant disorder | 5.0 (1.6)  | --          |
| Conduct disorder              | 2.2 (1.9)  | --          |
| DBDRS symptom counts          |            |             |
| Hyperactivity/impulsivity     | 5.5 (2.4)  | 5.5 (3.0)   |
| Inattention                   | 7.6 (2.5)  | 6.4 (3.3)   |
| Oppositional defiant disorder | 6.3 (2.3)  | 4.5 (3.2)   |
| Conduct disorder              | 2.3 (2.5)  | 1.0 (1.2)   |
| APSD total score              | 80.9 (7.1) | 74.5 (7.1)  |
| ICU total score               | 37.5 (5.0) | 40.9 (12.4) |
| CPS-R total score             | 35.9 (5.0) | 31.8 (10.1) |

*Note.* Values in table are means with standard deviations in parentheses. DBDSPI = Disruptive Behavior Disorders Structured Parent Interview; DBDRS = Disruptive Behavior Disorders Rating Scale; ICU = Inventory of Callous-Unemotional Traits; APSD = Antisocial Process Screening Device; CPS-R = Child Psychopathy Scale–Revised.

the Disruptive Behavior Disorders Rating Scale (DBDRS; Pelham, Gnagy, Greenslade, & Milich, 1992) and the Impairment Rating Scale (IRS; Fabiano et al., 2006) and parent report on the Disruptive Behavior Disorders Structured Parent Interview (DBDSPI; Pelham, Greiner, & Gnagy, 2006). Following DSM-IV criteria (American Psychiatric Association, 2000), diagnoses were assigned if a sufficient number of symptoms were endorsed and if the child evidenced clinically significant impairment. Of the 11 participants, 6 (55%) were diagnosed with ODD/ADHD and 5 (45%) were diagnosed with CD/ADHD. Rating scale data for the participants are summarized in Table 1.

Children also were selected based on the presence of elevated levels of CU traits. Our primary measure of CU traits was parent and teacher ratings on the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). This measure was selected because it is currently the only measure of CU traits with established psychometric properties, including normative data, to assess elementary school-age children. Norms were developed for the APSD using a large community sample of school-age children recruited from referrals to mental health clinics (Frick & Hare, 2001). All participants had *T* scores at or above 65 on the CU scale as rated by parent, teacher, or both, which indicates a clinically significant level. We also measured CU traits using the Inventory of Callous-Unemotional Traits (ICU; Frick, 2004) and the Child Psychopathy Scale–Revised (CPS-R; Lynam, 1997), which have been

used in several studies but do not yet have published norms available. Means and standard deviations for these measures also are reported in Table 1.

## PROCEDURE

### Overview

All procedures used in this study were approved by a university institutional review board and parent consent and child assent were obtained prior to initiating the research. Children were enrolled in the 2011 Summer Treatment Program (STP) conducted in the southeast United States. The STP is an 8-week intensive treatment program for children with clinically significant attention or behavior problems (Pelham, Greiner, & Gnagy, 1998). The STP was conducted from 8:00 a.m. until 5:00 p.m. on weekdays. During the STP day children participated in the following activities: two academic classes and an art class, three group recreational activities, swimming, lunch, and recess. Children were placed in a single group of 12 children who remained together throughout treatment. A lead counselor supervisor and four undergraduate counselors (outside of the classroom), and a teacher and teacher aide (in the classroom) implemented treatment. During recreational periods, children played soccer, softball, or basketball, and during classroom periods children worked on individualized academic assignments. A more extensive description of the STP is available elsewhere (Pelham et al., 2010).

The first week of the STP was considered a warm-up week, in which both the counselors and children were learning procedures; data from this week were not used in this study. After the first week, treatment was implemented in 1-week blocks for 7 consecutive weeks. Treatment was evaluated in a group setting using an alternating treatment manipulation that followed an A-B-A-C-A-BC-A design, with conditions designated as follows: A = standard treatment, B = low punishment treatment, C = high-reward treatment, and BC = combined treatment (low punishment and high reward). The experimental design follows earlier research that used reversal designs to examine treatment effects in small-group settings (Ayllon & Azrin, 1967). Order of the nonstandard treatments (B, C, BC) was determined randomly. Prior to treatment, counselors were intensively trained and tested on STP procedures (see Pelham et al., 2010, for details). During treatment, counselors were observed for at least 1 hour every day by clinical supervisors and given feedback on treatment implementation to maintain treatment integrity. To facilitate the transition between treatment conditions, counselors had weekly meetings with clinical supervisors to prepare for and practice changes to treatment protocol, and new procedures

were explained and discussed with children at the onset of each new condition. Counselors were blinded to children's conduct problem diagnoses and CU levels, but were not blinded to treatment condition as they were implementing it. Data analyses also were conducted by individuals who were not blinded to treatment condition.

#### *Standard Treatment*

In standard treatment, children's behavior was shaped through daily and weekly reinforcers for appropriate behaviors, and daily and weekly punishment for inappropriate behaviors. The primary daily reinforcement was the chance to participate in recess at lunch and again at the end of the day. The primary weekly reinforcer was the chance to participate in "Fun Friday," which was a different fun activity each week (video game time, swim time, etc.). Punishers included detention in place of recess, during which children were required to write sentences, and detention in place of a Fun Friday, during which children were required to complete chores (cleaning desks, sweeping floors, etc.). In addition, a 10-minute time-out was assigned for intentional aggression, intentional destruction of property, and repeated noncompliance. Whether children earned reward, punishment, or neither (a neutral outcome), each day and week was based on their performance on (a) a daily behavior report card (DRC); and (b) the comprehensive token economy that awarded children points for positive and nonnegative behaviors, and took points away for negative behaviors. The DRC consisted of a list of treatment goals that were individualized to each child and were evaluated throughout the day. Children earned daily rewards if they met a high percentage (>80%) of their treatment goals, or received punishment if they met a low percentage (<50%) of their treatment goals. Parents were instructed to continue with a normal evening routine (i.e., no reward or punishment) when children met a moderate percentage (50–80%) of treatment goals. Weekly consequences (reward, punishment, or neutral) were based on DRC performance and on the number of points they earned during the week. Parents were given feedback about their child's behavior at the end of each day and instructed to deliver the same daily or weekly outcome (reward, punishment, neutral) at home that their child received at the STP for that day or week.

#### *Low-Punishment Treatment*

In the low-punishment condition, the same procedures were followed as in standard treatment but with the following modifications. First, time-out was reduced from 10 minutes to 2 minutes. Second, daily and weekly punishments were dropped such that daily and weekly consequences consisted of two levels (reward

or neutral). That is, children either earned recess as a daily reward or they did not (i.e., children sat out of recess but sentence writing during recess was discontinued as a punishment) and children either earned Fun Friday as a weekly reward or they did not (i.e., children participated in a typical STP day and chores were discontinued as a punishment). Third, children did not lose points for inappropriate behavior, although we continued to identify inappropriate behavior (e.g., "you're interrupting me" rather than "you lose 20 points for interruption"), and DRCs were revised to include only positive treatment goals (e.g., contributing to group discussions or participating during activity periods). At home parents were encouraged to either reward children for achieving most of their DRC goals (>80% goals met), or continue with the normal evening routine (<80% goals met).

#### *High-Reward Treatment*

The high-reward condition matched the standard condition except that children had the opportunity to earn material rewards in addition to fun activities. Material rewards (toys) were earned through a ticket system in which children were awarded tickets every 15 minutes that they did not exhibit negative behavior. They also were awarded tickets for exhibiting positive behaviors. Children exchanged their tickets for toys at the end of the week, with better toys requiring more tickets. Toys were identified using a reinforcement interview, which was conducted with parents and children. Children were given frequent reminders of the reinforcement associated with tickets both verbally (during group discussion and individual conversations with counselors) and by allowing them to visit the toy store multiple times per day.

#### *Combined Treatment*

The combined treatment condition consisted of simultaneously implementing the procedures used in the low-punishment and high-reward treatments. Conceptually, children were given high incentive not to misbehave while simultaneously given as little attention as feasible for misbehavior.

#### DIAGNOSTIC MEASURES

##### *Disruptive Behavior Disorders Rating Scale*

The DBDRS consists of 45 questions designed to measure DSM-IV symptoms of ADHD, ODD and CD using a 4-point Likert scale (0 = *not at all* to 3 = *very much*; Pelham et al., 1992). Validity of the DBDRS has been supported by research showing significant convergent and discriminant validity as compared with other measures of psychopathology (Massetti, Pelham, & Gnagy, 2005) and as compared

with observations of children's behavior (Wright, Waschbusch, & Frankland, 2007). The internal consistency (Cronbach's alpha) for our sample ranged from .78 to .90 and .86 to .95 across parent- and teacher-reported symptom scales, respectively.

#### *Disruptive Behavior Disorders Structured Parent Interview*

The DBDSPSI is a structured interview designed to measure DSM-IV symptoms of ADHD, ODD, and CD (Pelham et al., 2006). The DBDSPSI was administered by trained interviewees (Ph.D.-level clinicians or advanced graduate students) to parents in a one-on-one setting. Clinicians introduced each symptom by asking parents to describe their child's typical behavior over the past year using situational prompts (e.g., at home, at school) provided in the interview. Based on parental responses, clinicians rated how much the symptom was a problem, in each situation and overall, using operationally defined scales that ranged from "not a problem" to "severe problem." Symptoms were counted as present if the overall rating indicated it was a moderate or severe problem within the last 6 months. The internal consistency ranged from .84 to .93 (Cronbach's alpha) across symptom subscales.

#### *Antisocial Process Screening Device*

The APSD consists of 20 questions designed to measure CU traits, narcissism, and impulsivity (Frick & Hare, 2001). Only the six items from the CU scale were used in the current study, and these were rated on a 3-point Likert scale (0 = *not at all true* to 2 = *definitely true*). The validity of the CU scale on the APSD has been supported in numerous studies (Kotler & McMahon, 2010). Internal consistency within our sample was consistent with past research (Cronbach's alpha = .56 and .50 for parent and teacher report, respectively; e.g., Bijttebier & Dacoene, 2009), however, these values are likely an underestimate of internal consistency because of our small sample size.

#### *Inventory of Callous-Unemotional Traits*

The ICU consists of 24 items designed to measure the extent to which children are uncaring, callous, and unemotional (Frick, 2004). Items were rated using a 4-point Likert scale (0 = *not at all true* to 3 = *definitely true*). Although norms for the ICU have not yet been established for elementary school-age samples, preliminary research has supported its validity and reliability in this age range (Houghton, Hunter, & Crow, 2012). The internal consistency for this measure in our sample was good for parent and teacher report (Cronbach's alpha = .82 and .89, respectively).

#### *Child Psychopathy Scale–Revised*

The CPS-R consists of 55 items designed to measure numerous facets of psychopathic personality, including callous-unemotional traits (Lynam, 1997). Items were rated as present or absent using a yes/no response format and, after reverse coding appropriate items, summed into a total score. Although norms for the CPS-R are not available for elementary school-age samples, preliminary research has supported the reliability and validity of the total score in youth as young as 12 years old (Kotler & McMahon, 2010). This measure had good internal consistency for our sample (parent-report alpha = .73, teacher-report alpha = .82).

#### DEPENDENT MEASURES

##### *Overview*

The dependent measures were frequency counts of participant behavior as recorded by counselors throughout the day in the context of implementing the comprehensive token economy (i.e., a point system). The current study used data from the three sports activity periods and from time-out. We excluded data from transitions (i.e., walking in line from one location to another), lunch, and swimming because treatment was most consistently implemented during the sports periods (i.e., children spent a considerable portion of transitions and swimming in locker rooms and bathrooms, where it was difficult to accurately monitor them, and rates of behavior during lunch were low, when they were occupied by eating). Point-system data were not collected during the three classroom periods, as another form of behavioral management (response-cost) was used in the classroom (for details refer to the STP manual; Pelham et al., 1998). Finally, this study used data from Monday through Thursday because children spent Friday afternoon receiving their weekly contingency (reward, neutral, or punishment activity, depending on their behavior and treatment condition), which necessitated a nonstandard schedule on Fridays.

##### *Negative Behavior During Sport Activities*

Negative behavior during sport activities was operationalized using the following point-system behavioral categories: aggression, destruction of property, noncompliance, stealing, lying, verbal abuse to staff, teasing peers, and swearing (see Pelham et al., 1998, for operational definitions). These behaviors were combined in a single negative behavior score for parsimony. Frequency counts of these behaviors were summed across the three sports periods for each day. If a child was missing for a portion of the day, his or her daily behavior totals were prorated based on the time he or she was present. Using these totals, we calculated a daily average of negative behavior for

each week of treatment for each child. Past research using STP point-system data has found high inter-rater reliability and strong correlations between point-system categories and other measures of behavioral problems, such as parent report or clinical interview (Wright et al., 2007).

#### *Negative Behavior During Time-Out*

As noted earlier, time-out was assigned following every instance of intentional aggression, intentional destruction of property, or repeated noncompliance. Negative behavior was measured using the same definition described earlier (a sum of the frequency counts of lying, stealing, destruction of property, etc.) but applying this definition to times the child was in time-out. However, because the duration of time-out was manipulated as a function of treatment condition, we did not use a total frequency count of negative behavior (nor did we use duration or frequency of time-out) as the dependent variable. Instead, we used the rate of negative behavior during time-out, computed by dividing the frequency of negative behavior during time-out by the total minutes of time-out. This was computed each day for each participant.

#### DATA ANALYSES

For both sport activity and time-out data, we followed earlier research (Ayllon & Azrin, 1967) by first examining group-level treatment effects and then examining individual differences in treatment response. Group effects were examined using repeated measures ANOVAs with treatment condition as the independent variable. For the purpose of the ANOVAs, standard treatment was separated into two treatment conditions: the mean of negative behavior during the first 2 weeks of standard treatment (standard 12), and the mean of negative behavior during the second 2 standard weeks (standard 34). This was justified because negative behavior was appreciably lower during the first 2 weeks of standard treatment compared with the latter 2 standard weeks on both activity and time-out measures,  $t(10) = -2.83$ ,  $p = .02$ , and  $t(8) = -2.71$ ,  $p = .03$ , respectively. (see Table 2) Thus, the ANOVA compared the following

conditions: standard12 versus standard34 versus low punishment versus high reward versus combined treatment. Assumption testing was conducted with Mauchley's test of sphericity, and analyses violating this assumption were corrected using the Greenhouse-Geiser degrees of freedom adjustment. Significant differences were followed up with multiple comparisons using  $t$  tests based on a least significant difference. After examining these group effects, individual differences in treatment response were explored by plotting data separately for each participant. All analyses were conducted using SPSS (version 20.0).

## Results

### SPORTS ACTIVITIES

#### *Group Effects*

Using repeated measures one-way ANOVA, there was a significant main effect of condition on levels of negative behavior during activities,  $F(2, 16) = 5.28$ ,  $p = .02$ ,  $\eta_p^2 = .37$ . The assumption of sphericity was violated, therefore, degrees of freedom were adjusted. Examination of means (see Table 2) and post hoc multiple comparisons showed that negative behavior was highest during the combined condition, with rates of negative behavior that were significantly higher than any other condition ( $ps < .03$ ; Hedges's  $g$  for combined condition with standard12 = .69, low punishment = .92, high reward = .92, standard 34 = .41). The standard34 and high-reward conditions followed, and the standard34 condition had significantly higher rates of negative behavior than the standard12 condition ( $p = .01$ , Hedges's  $g = .33$ ), and marginally significantly higher than the low-punishment condition ( $p = .09$ , Hedges's  $g = .63$ ). However, levels of negative behavior during the high-reward condition were not significantly different from any other condition excluding the combined condition. Finally, negative behavior was lowest during the low-punishment and standard12 conditions, which did not significantly differ from each other.

#### *Individual Differences*

Descriptive analyses of individual differences in response to treatment (see Figure. 1) suggested a number of interesting patterns. First, a handful of

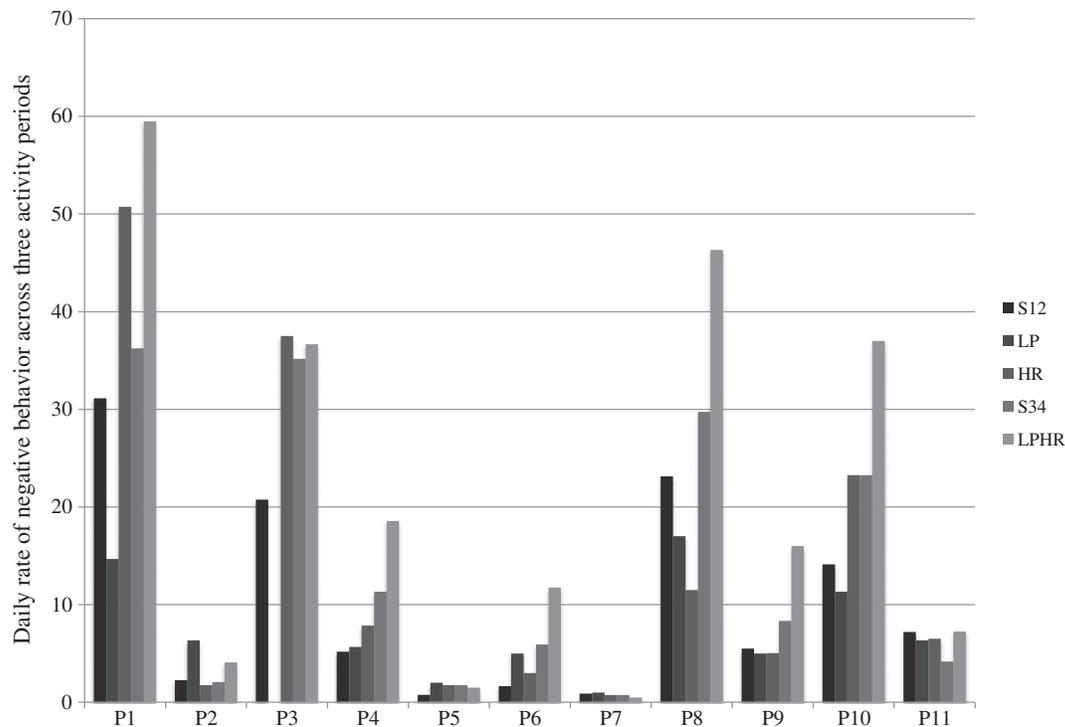
Table 2

Average Daily Rate of Negative Behavior During Three Activity Periods and per Minute of Time-Out

|            | S12                        | LP                        | HR                          | S34                        | LPHR                       | F-value |
|------------|----------------------------|---------------------------|-----------------------------|----------------------------|----------------------------|---------|
| Activities | 10.23 <sup>a</sup> (10.49) | 7.43 <sup>ac</sup> (5.24) | 13.60 <sup>ac</sup> (16.63) | 14.43 <sup>c</sup> (13.94) | 21.74 <sup>b</sup> (20.03) | 5.28*   |
| Time-out   | 0.09 <sup>c</sup> (0.17)   | 0.05 <sup>ac</sup> (0.11) | 0.30 <sup>ad</sup> (0.24)   | 0.51 <sup>bd</sup> (0.44)  | 0.68 <sup>b</sup> (0.57)   | 8.68*** |

Note. S12 = standard12; LP = low punishment; HR = high reward; S34 = standard34; LPHR = combined. Means with differing subscripts are significantly different at  $p < .05$  based on least-significant difference post hoc comparisons.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .



**FIGURE 1** Average daily rate of negative behavior during activities for each participant. Note. LP = low punishment, HR = high reward, LPHR = combined, S12 = standard12, and S34 = standard34.

children demonstrated a similar response pattern to the group-level pattern, albeit at varying levels of negative behavior severity (i.e., Participants 1, 4, and 10). On the other hand, some children responded well to both low punishment and high reward in comparison with the standard treatment conditions (i.e., Participants 6, 8, and 9). Finally, some children demonstrated consistently low levels of negative behavior throughout the summer (i.e., Participants 2, 5, 7, and 11), whereas others demonstrated consistently high levels of negative behavior throughout the summer (e.g., Participant 3 [absent during low punishment due to physical injury that occurred outside of treatment]).

#### TIME-OUT

##### *Group Effects*

Repeated measures one-way ANOVA revealed a significant main effect of condition on levels of negative behavior,  $F(4, 20) = 8.68$ ,  $p < .001$ ,  $\eta_p^2 = .64$ . Examination of means (see Table 2) and post hoc multiple comparisons showed that time-out data largely paralleled the sports activities data. First, the combined condition had significantly higher rates of negative behavior compared with any other condition ( $ps < .04$ ; Hedges's  $g$  for combined

condition with standard12 = 1.35, low punishment = 1.44, high reward = .84), excluding standard34. The second-highest rates of negative behavior occurred during the standard34 condition. Negative behavior during the standard34 condition was significantly higher than the standard12 and low-punishment conditions ( $ps < .02$ ; Hedges's  $g$  for standard 34 with: standard12 = 1.21, low punishment = 1.35), and marginally significantly higher than the high-reward condition ( $p = .09$ , Hedges's  $g = .57$ ). Negative behavior during the high-reward condition was significantly higher than during the standard12 condition ( $p = .04$ , Hedges's  $g = .97$ ), but did not significantly differ from rates of negative behavior during the low-punishment condition. Finally, the lowest rates of negative behavior occurred during the standard12 and low-punishment conditions, both of which were close to zero and did not significantly differ.

##### *Individual differences*

Individual differences during time-out also paralleled individual differences during sports activities (see Figure 2). Again, Participants 1 and 10 demonstrated a similar response pattern as group-level performance. However, Participant 4 demonstrated relatively

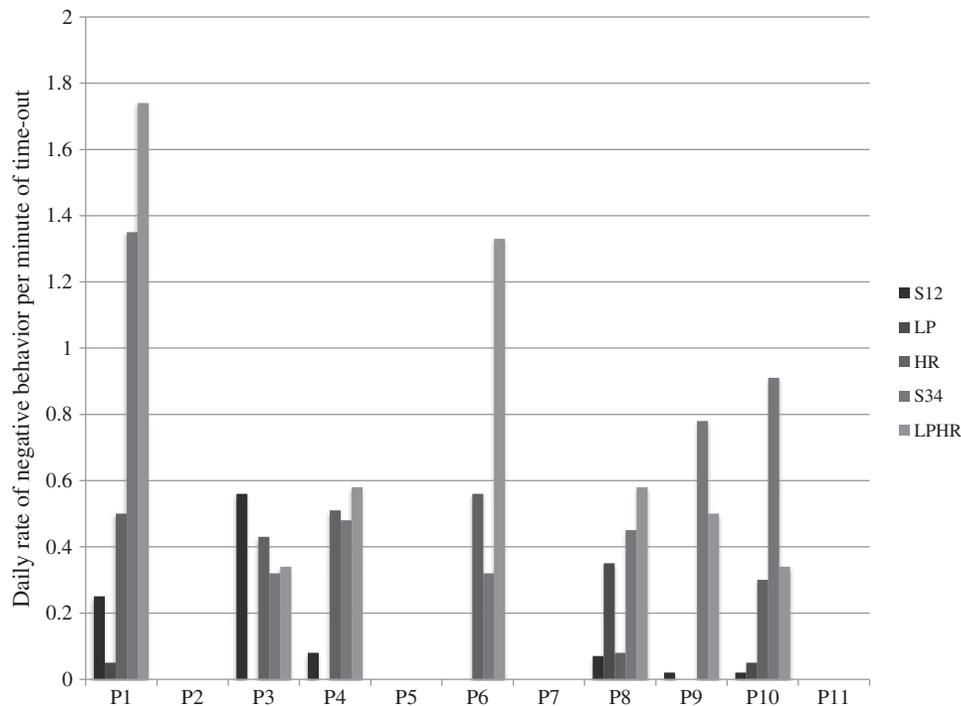


FIGURE. 2 Average daily rate of negative behavior per minute of time-out for each participant.

uniform and high levels of negative behavior across the high-reward, combined, and standard34 conditions. Further similarities between time-out and activity performance were noted among children with low base rates of negative behavior during activities, who were rarely in time-out, and never demonstrated any negative behaviors at any point during their time-outs. Interestingly, there was no consistent response pattern within time-out among children who responded well to both low-punishment and high-reward conditions during activities (i.e., Participants 6, 8, and 9). Only Participant 9 replicated his or her activity response pattern during time-out, whereas Participant 6 performed relatively poorly during the high-reward condition and Participant 8 performed relatively poorly during the low-punishment condition on time-out measures. Finally, Participant 3 maintained relatively high levels of negative behavior during time-out across all conditions for which this child was present.

### Discussion

The current study examined the treatment response of 11 children with CPCU to four modifications of behavioral therapy. Group-level analyses indicated the best treatment response occurred during the low-punishment and standard12 conditions, whereas the worst response occurred during the

combined condition. Frequencies of negative behaviors during the high-reward and standard34 conditions fell between these points. These results were similar across both activity and time-out measures. Thus, our hypothesis that children would demonstrate less negative behavior during the modified treatment conditions relative to the standard treatment conditions was partially supported in that negative behavior was lowest during the low-punishment condition relative to the latter standard treatment condition. In contrast, we did not find evidence in support of the efficacy of either the high-reward or combined treatment conditions, as levels of negative behavior during both the high-reward and combined conditions were equal to or greater than levels of negative behavior during standard treatment. Unexpectedly, the highest rates of negative behavior actually occurred during the combined condition.

These results support the findings of previous research suggesting that children with CU traits are less sensitive to punishment while engaging in goal-directed behavior (Dadds & Salmon, 2003), and consequently less likely to change their behavior in the face of punishment. First, excluding the combined condition, we found relatively higher rates of negative behavior both during activities and in time-out when standard punishment methods

were used (i.e., high-reward and standard conditions). Also, negative behavior during the standard conditions worsened over time, suggesting that punishment strategies were doing little to ameliorate negative behavior. In contrast, we observed the lowest rates of negative behavior during activities and time-out in the low-punishment condition. Indeed, none of the children demonstrated *any* aggression, destruction of property, stealing, or lying during time-out in the low-punishment condition. This is a promising finding, given that previous research has found that time-out is particularly difficult to implement among children with CU traits (Haas et al., 2011). It is possible that standard punishment methods (e.g., point loss, 10-minute time-out, detention during recess and Fun Friday) attenuated children's motivation to behave appropriately as punishment interfered with their ability to earn rewards. In the face of seemingly impossible goals, other types of reinforcers, such as peer attention or counselor negative attention, may have become more salient and enticing. Consequently, children may have engaged in higher rates of negative behavior as a means to these ends. Furthermore, the increasing rates of negative behavior during the standard conditions also are interesting when compared with past STP research of children with ADHD only, which report significant reductions in negative behavior over 8 weeks of standard treatment (i.e., standard condition; Pelham et al., 2000). Therefore, our results suggest that worsening treatment response may be an effect of CU traits—although this is speculative as we did not have a comparison ADHD-only group.

These interpretations also are tempered by the wide variability of response patterns across participants. Indeed, three response patterns were evident in our inspection of the individual treatment response graphs. Parallel to the group analysis, some children appeared to respond best to the low-punishment and standard conditions; in contrast, other children seemed to respond best to the high-reward condition. Finally, a handful of children demonstrated rates of negative behavior similar to those of normative samples across all treatment conditions (Pelham et al., 1998). Although all children in our sample shared the same clinical profiles (i.e., CPCU with ADHD), it is curious how variable their treatment responses were. It is likely that other environmental and child characteristics moderated each child's response pattern. For instance, in our opinion children who responded well were more motivated than others to bond with counselors, consequently child-counselor rapport may have been an important factor in treatment responsiveness. Further, some children demonstrated greater self-control and presumably were able to manage urges to misbehave

when these urges ran counter to their goals. On the other hand, other children reacted impulsively to challenging situations (e.g., trying to sit in classroom quietly while peers were being disruptive), meaning they had fewer opportunities to be reinforced/rewarded for appropriate behavior, and more opportunities to be reinforced for inappropriate behavior via peer attention (e.g., other children cheer when child runs around classroom). Obviously these suggestions are speculative, and future research is necessary to understand the wide variability of treatment responsiveness of children with CPCU.

In spite of this heterogeneity, two consistencies were noted: (a) the combined condition had the highest rates of negative behavior, and (b) there was a trend for children's behavior to worsen over time (i.e., rates of negative behavior during the first 2 weeks of standard treatment were significantly less than the latter 2 weeks on most behavioral measures). Given these consistencies, it is likely that ordering effects influenced treatment response patterns. It may be that the low-punishment condition outperformed the other conditions because it appeared earlier in the STP (during the third week), whereas the combined treatment may have underperformed relative to other conditions because it occurred later in the STP (during the seventh week). That is, even without manipulating treatment children's behavior may have differed earlier versus later in treatment as peers and counselors become more familiar with each other and as peers and counselors become more fatigued by the demanding nature of the treatment. These and other factors cannot be disentangled from the treatment effects in this study. To do so, order of treatment would have to be counterbalanced across participants, but the small sample size used in this pilot study did not allow for this possibility.

Related to this point, there is anecdotal evidence that experience within each treatment modification played an important role in children's treatment response. Children appeared to show a more positive response to treatment procedures when they were first introduced as compared with when they had gained some experience with them. For instance, the lowest rates of negative behavior occurred the first day the children were exposed to new treatment procedures. Interestingly, this pattern did not hold for across the standard treatment weeks, suggesting it was the novelty of a new system rather than the start of a new week that influenced treatment response. Consistent with this observation, children's behavior was often at its worst when they had experience with treatment procedures and gained knowledge of how to exploit those procedures. For instance, during the combined treatment condition counselors were not allowed to take points for negative behaviors but they

were allowed to give commands to stop exhibiting negative behaviors for 2 minutes when children were being so disruptive that the group could not function. Because children earned points for compliance when they followed this command, they quickly learned to exhibit high rates of negative behavior so their counselors would issue a command for them to stop, which in turn earned them points for compliance. Thus, the poor response to treatment in the combined condition was likely due to the fact that high rates of negative behavior had no appreciable negative consequence (point loss) but did have relatively predictable positive consequences (points for complying with a command to stop misbehaving). More generally, these observations suggest that novelty of treatment may be especially important to consider in designing treatments for children with CU traits, and we suspect that boredom and familiarity with treatment protocols (and their weaknesses) contributed to increasing negative behavior.

Deviancy training, in which children's misbehavior is reinforced by the social reactions of peers, is thought to be especially problematic in group-based treatments for children with antisocial behavior (Dishion & Tipsord, 2011). This may have been evident in the present study as well, in that we observed one child frequently encourage other children to behave aggressively toward counselors while they were in time-out. This appeared to occur most often in the combined treatment condition that inadvertently delivered the least-effective treatment. If this observation is confirmed in future research, it would suggest that deviancy training may be important to consider when treating children with CU traits in group settings.

This study was the first (that we are aware of) to modify behavioral treatment procedures for use with children with CPCU in a group-based format. Although this is a novel study, results should be interpreted with consideration of the limitations of this study. First, as previously mentioned, treatment conditions were presented in a single order, making it impossible to parse ordering effects and practice/time effects from treatment effects. Second, our sample was small, and it is possible that we did not have the power to detect certain effects. Further, because of the small sample size and the need to take Type II error into account, we did not correct for family-wise error in our post hoc testing and our comparisons are vulnerable to Type I error. It also was impossible to control for participant characteristics that may have moderated treatment response, such as intelligence, age, or CU trait severity. Nonetheless, we tried to limit these effects by restricting the range of these variables (e.g., recruiting children with clinically significant CU traits only, limiting age range to

school-age children). Third, this data did not control for counselor behavior. Although the counselors were implementing a very structured treatment and were routinely observed, it is possible that their behavior toward the children changed as they developed relationships with the children. For example, one counselor became a favorite of the children and over the course of the summer it became increasingly difficult for her to implement time-out effectively because the children found any time with her to be reinforcing. Fourth, the DRC goals were changed weekly and may have accounted for some variance in treatment response, and we were unable to formally assess parents' implementation of these goals. Also, the weekly changes in treatment condition may have been difficult for children to adjust to and longer treatment periods would have been more desirable. Finally, our treatment manipulations were behaviorally based and did not focus on the relationships among children or between children and staff. Recent research suggests that children with CU traits tend to have an insecure attachment style (Pasalich, Dadds, Hawes, & Brennan, 2012) and consequently may benefit from emotion-focused interactions with positive role models.

Future research should address these limitations in several ways. Of most importance, these results should be replicated in studies using multiple treatment groups that randomly assign children to treatment and that counterbalance the order of treatment conditions. Incorporating both behavioral and attachment perspectives into a treatment model may improve treatment response and possibly reduce deviancy training effects. Future research also should examine how STP procedures could be translated into parent-child treatment models, as it is a more accessible form of psychological treatment.

#### Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

#### References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev. ed.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ayllon, T., & Azrin, N. H. (1967). The measurement and reinforcement of behavior of psychotics. *Journal of the Experimental Analysis of Behavior*, 8, 357-383.
- Bijttebier, P., & Dacoene, S. (2009). Assessment of psychopathic traits in children and adolescents: Further validation of the Antisocial Process Screening Device and the Childhood Psychopathy Scale. *European Journal of Psychological Assessment*, 25, 157-163. <http://dx.doi.org/10.1027/1015-5759.25.3.157>

- Blair, R. J. R., Colledge, E., & Mitchell, D. G. V. (2001). Somatic markers and response reversal: Is there orbito-frontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology*, 29, 499–511. <http://dx.doi.org/10.1023/a:1012277125119>
- Conduct Problems Prevention Research Group. (2011). The effects of the Fast Track preventive intervention on the development of conduct disorder across childhood. *Child Development*, 82, 331–345. <http://dx.doi.org/10.1111/j.1467-8624.2010.01558.x>
- Dadds, M. R., & Salmon, K. (2003). Punishment insensitivity and parenting: Temperament and learning as interacting risks for antisocial behavior. *Clinical Child and Family Psychology Review*, 6, 69–86. <http://dx.doi.org/10.1023/a:1023762009877>
- Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social development. *Annual Review of Psychology*, 62, 189–214. <http://dx.doi.org/10.1146/annurev.psych.093008.100412>
- 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 & Adolescent Psychology*, 37, 215–237.
- Fabiano, G. A., Pelham, J. W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., . . . Burrows-Maclean, 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 & Adolescent Psychology*, 35, 369–385. [http://dx.doi.org/10.1207/s15374424jccp3503\\_3](http://dx.doi.org/10.1207/s15374424jccp3503_3)
- Finger, E. C., Marsh, A. A., Mitchell, D. G., Reid, M. E., Sims, C., Budhani, S., . . . Blair, J. R. (2008). Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Archives of General Psychiatry*, 65, 586–594. <http://dx.doi.org/10.1001/archpsyc.65.5.586>
- Frick, P. J. (2004). *Inventory of callous-unemotional traits*. New Orleans, LA: University of New Orleans.
- Frick, P. J., Cornell, A. H., Bodin, S. D., Dane, H. E., Barry, C. T., & Loney, B. R. (2003). Callous-unemotional traits and developmental pathways to severe conduct problems. *Developmental Psychology*, 39, 246–260. <http://dx.doi.org/10.1037/0012-1649.39.2.246>
- Frick, P. J., & Ellis, M. (1999). Callous-unemotional traits and subtypes of conduct disorder. *Clinical Child and Family Psychology Review*, 2, 149–168. <http://dx.doi.org/10.1023/a:1021803005547>
- Frick, P. J., & Hare, R. D. (2001). *The Antisocial Process Screening Device (APSD)*. Toronto, Canada: Multi-Health Systems.
- 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. <http://dx.doi.org/10.1007/s10802-010-9480-4>
- Hawes, D. J., & Dadds, M. R. (2005). The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting and Clinical Psychology*, 73, 737–741. <http://dx.doi.org/10.1037/0022-006x.73.4.737>
- Houghton, S., Hunter, S. C., & Crow, J. (2012). Assessing callous-unemotional traits in children aged 7- to 12-years: A confirmatory factor analysis of the Inventory of Callous-Unemotional Traits. *Journal of Psychopathology and Behavioral Assessment*, 35, 215–222. <http://dx.doi.org/10.1007/s10862-012-9324-3>
- Kimonis, E. R., & Armstrong, K. (2012). Adapting parent-child interaction therapy to treat severe conduct problems with callous-unemotional traits: A case study. *Clinical Case Studies*, 11, 234–252.
- Kotler, J. S., & McMahon, R. J. (2010). Assessment of child and adolescent psychopathy. In R. T. Salekin & D. T. Lynam (Eds.), *Handbook of child and adolescent psychopathy* (pp. 79–109). New York, NY: Guilford Press.
- Lynam, D. R. (1997). Pursuing the psychopath: Capturing the fledgling psychopath in a nomological net. *Journal of Abnormal Psychology*, 106, 425–438. <http://dx.doi.org/10.1037/0021-843x.106.3.425>
- Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, 116, 155–165. <http://dx.doi.org/10.1037/0021-843x.116.1.155>
- Masi, G., Manfredi, A., Milone, A., Muratori, P., Polidori, L., Ruglioni, L., & Muratori, F. (2011). Predictors of non-response to psychosocial treatment in children and adolescents with disruptive behavior disorder. *Journal of Child and Adolescent Psychopharmacology*, 21, 51–55. <http://dx.doi.org/10.1089/cap.2010.0039>
- Massetti, G. M., Pelham, W. E., & Gnagy, E. M. (2005, June). *Situational variability of ADHD, ODD, and CD: Psychometric properties of the DBD interview and rating scale for an ADHD sample*. Paper presented at the annual meeting of the International Society for Research in Child and Adolescent Psychopathology, New York, NY.
- McMahon, R. J., Witkiewitz, K., & Kotler, J. S. (2010). Predictive validity of callous-unemotional traits measured in early adolescence with respect to multiple antisocial outcomes. *Journal of Abnormal Psychology*, 119, 752–763. <http://dx.doi.org/10.1037/a0020796>
- Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, 8, 399–424. <http://dx.doi.org/10.1017/S0954579400007161>
- Nock, M. K., Kazdin, A. E., Hiripi, E., & Kessler, R. C. (2006). Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the National Comorbidity Survey Replication. *Psychological Medicine*, 36, 699–710. <http://dx.doi.org/10.1017/s0033291706007082>
- Nock, M. K., Kazdin, A. E., Hiripi, E., & Kessler, R. C. (2007). Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: Results from the National Comorbidity Survey Replication. *Journal of Child Psychology and Psychiatry*, 48, 703–713. <http://dx.doi.org/10.1111/j.1469-7610.2007.01733.x>
- O'Brien, B. S., & Frick, P. J. (1996). Reward dominance: Associations with anxiety, conduct problems, and psychopathy in children. *Journal of Abnormal Child Psychology*, 24, 223–240. <http://dx.doi.org/10.1007/bf01441486>
- Pasalich, D. S., Dadds, M. R., Hawes, D. J., & Brennan, J. (2012). Attachment and callous-unemotional traits in children with early-onset conduct problems. *Journal of Child Psychology and Psychiatry*, 53, 838–845. <http://dx.doi.org/10.1111/j.1469-7610.2012.02544.x>
- 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 & Adolescent Psychiatry*, 31, 210–218.
- 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. <http://dx.doi.org/10.1023/A:1005127030251>

- Pelham, W. E., Jr., Gnagy, E. M., Greiner, A. R., Waschbusch, D. A., Fabiano, G. A., & Burrows-Maclean, L. (2010). Summer treatment programs for attention-deficit/hyperactivity disorder. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed., pp. 277–292). New York, NY: Guilford Press.
- Pelham, W. E., Jr., Greiner, A. R., & Gnagy, E. M. (1998). *Children's summer treatment program manual*. Buffalo, NY: Comprehensive Treatment for Attention Deficit Disorders.
- Pelham, W. E., Jr., Greiner, A. R., & Gnagy, E. M. (2006). *Disruptive behavior disorders clinical interview*. Buffalo, NY: Comprehensive Treatment for Attention Deficit Disorders.
- Robins, L. N., & Price, R. K. (1991). Adult disorders predicted by childhood conduct problems: Results from the NIMH Epidemiologic Catchment Area project. *Psychiatry: Interpersonal and Biological Processes*, *54*, 116–132.
- Waschbusch, D. A., Walsh, T. M., Andrade, B. F., King, S., & Carrey, N. J. (2007). Social problem solving, conduct problems, and callous-unemotional traits in children. *Child Psychiatry and Human Development*, *37*, 293–305. <http://dx.doi.org/10.1007/s10578-006-0033-6>
- Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology*, *65*, 93–109. <http://dx.doi.org/10.1037/0022-006x.65.1.93>
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *71*, 136–149. <http://dx.doi.org/10.1037/0022-006x.71.1.136>
- Wright, K. D., Waschbusch, D. A., & Frankland, B. W. (2007). Combining data from parent ratings and parent interview when assessing ADHD. *Journal of Psychopathology and Behavioral Assessment*, *29*, 141–148. <http://dx.doi.org/10.1007/s10862-006-9039-4>

RECEIVED: July 12, 2013

ACCEPTED: October 21, 2013

Available online xxxx