



Parent Training Adapted to the Needs of Children With Callous–Unemotional Traits: A Randomized Controlled Trial

Georgette E. Fleming

Bryan Neo

Nancy E. Briggs

Silvana Kaouar

University of New South Wales

Paul J. Frick

Louisiana State University

Eva R. Kimonis

University of New South Wales and Karitane

Callous–unemotional (CU) traits designate a distinct subgroup of children with early-starting, stable, and aggressive conduct problems. Critically, traditional parenting interventions often fail to normalize conduct problems among this subgroup. The aim of this study was to test whether parent–child interaction therapy (PCIT) adapted to target distinct deficits associated with CU traits (PCIT-CU) produced superior outcomes relative to standard PCIT. In this proof-of-concept trial, 43 families with a 3- to 7-year-old child (M age = 4.84 years, SD = 1.12, 84% male) with clinically significant conduct problems and elevated CU traits were randomized to receive standard PCIT (n = 21) or PCIT-CU (n = 22) at an urban university-based research clinic. Families completed five assessments measuring child conduct problems, CU traits, and empathy. Parents in both conditions reported good treatment acceptability and significantly improved conduct problems and CU traits during active treatment, with no between-group differences. However, linear mixed-effects models showed treatment gains

in conduct problems deteriorated for children in standard PCIT relative to those in PCIT-CU during the 3-month follow-up period (d s = 0.4–0.7). PCIT-CU shows promise for sustaining improvements in conduct problems for young children with conduct problems and CU traits, but requires continued follow-up and refinement.

Keywords: conduct disorders; callous–unemotional traits; parent training; limited prosocial emotions; psychopathic traits

SERIOUS CONDUCT PROBLEMS (i.e., conduct disorders) are among the most common childhood mental problems, with a combined worldwide prevalence of 5.7% (Polanczyk et al., 2015). The risk factors for childhood conduct problems vary considerably across individuals, and effective intervention requires individualizing treatment to the unique needs of children on different developmental pathways (see Frick et al., 2014b, for a review). Diagnostic classification systems recognize the importance of this causal heterogeneity by including a specifier for the diagnosis of conduct disorders called with limited prosocial emotions (LPE). For example, to meet this criteria in the *Diagnostic and Statistical Manual of Mental Disorders—5th edition* (DSM-5; American Psychiatric Association, 2013), a child must show two or more of the following symptoms persistently (≥ 12 months) across multiple relationships/settings: (a)

This trial was registered with the Australian and New Zealand Clinical Trial Registry (ACTRN12616000280404). We have no conflicts of interest. E.R.K. received competitive funding from the University of New South Wales for this work.

Address correspondence to Eva R. Kimonis, Ph.D., School of Psychology, The University of New South Wales, Sydney, NSW 2052, Australia. e-mail: e.kimonis@unsw.edu.au.

0005-7894/© 2022 Association for Behavioral and Cognitive Therapies. Published by Elsevier Ltd.

lack of remorse or guilt, (b) callous–lack of empathy, (c) lack of concern about performance (at school, work, in other important activities), and (d) shallow/deficient affect. A similar specifier is included in the *International Statistical Classification of Diseases and Related Health Problems*—11th edition (ICD-11; [World Health Organization, 2019](#)). The addition of the LPE specifier was informed by decades of research supporting the clinical utility of callous–unemotional (CU) traits (i.e., lack of remorse, empathy, and concern over poor performance, and shallow/deficient affect).

CU traits designate a distinct subgroup of youth showing early starting, severe, persistent, and aggressive conduct problems that are less responsive to traditional psychosocial interventions, relative to children with conduct disorders without CU traits ([Frick et al., 2014b](#); [Hawes et al., 2014](#)). It is important to note that the term “traits” is not meant to suggest that these characteristics of the child are unchangeable and unaffected by the child’s context—that is, while there is evidence that CU traits are substantially heritable ([Moore et al., 2019](#)) and are moderately stable, even in early childhood ([Frick et al., 2014a](#)), there is also clear evidence that these traits are influenced by both genetic and environmental risk factors ([Hyde et al., 2016](#)) and can be changed by certain parenting practices ([Waller et al., 2015a](#)). However, the term “traits” is used to refer to the fact that, whereas some children occasionally show these behaviors as part of normal development, they are predictive of serious conduct problems when shown consistently across relationships and settings ([Frick & Kemp, 2021](#)).

While children with conduct disorder symptoms and co-occurring CU traits (CP + CU) show improvements in both their CU traits and conduct problems following treatment, they often begin and end treatment with more severe conduct problems that meet diagnostic threshold for the disorder, relative to children with conduct problems alone (CP-only; [Hawes et al., 2014](#)). For example, [Bjørnebekk and Kjøbli \(2017\)](#) found that observed callousness moderated treatment outcomes for children ($N = 323$, <12 years old, M age = 8.69 years, $SD = 2.14$, 73% boys) with conduct problems receiving Parent Management Training (PMT)—Oregon model. Higher callousness scores predicted greater posttreatment parent- and teacher-rated delinquency and aggression and less therapist-rated treatment success. PMT, which is a first-line treatment for conduct disorders, is thought to produce unequal outcomes across levels of CU traits because it does not address the distinct

familial and neurocognitive mechanisms underlying the conduct problems of children with CU traits—that is, it has been hypothesized that PMT’s theory-driven focus on improving the effectiveness and consistency of parental discipline to produce child behavior change is undermined by the temperamentally fearless and punishment-insensitive learning styles of children with CP + CU, who may experience core behavioral discipline strategies, such as time-out, as less aversive than CP-only children ([Barker et al., 2011](#); [Byrd et al., 2014](#); [Garcia et al., 2018](#); [Hawes & Dadds, 2005](#)). In contrast, PMT’s positive reinforcement strategies were rated by parents of clinic-referred children with conduct disorders as equally effective for reducing conduct problems across varying levels of child CU traits ([Hawes & Dadds, 2005](#)). These findings suggest that modifying traditional behavioral therapies for CP + CU children by emphasizing positive reinforcement over punishment may enhance their treatment outcomes.

The unique risk factors underpinning the ontogeny and maintenance of conduct problems for children with CU traits are relatively well understood ([Moul et al., 2018](#)), and constitute the treatment targets of novel and adapted treatments for this population. For example, supplementing PMT with parent–child emotion recognition training, to target emotion recognition deficits associated with CU traits, was superior to PMT alone in improving empathy and reducing conduct problems among 6- to 16-year-old children ($N = 195$, M age = 10.52 years, $SD = 2.64$, 76% boys) with emotional and behavioral problems when CU traits were high ([Dadds et al., 2012](#)). However, the positive impact of this enhancement on reducing conduct problems in children high on CU traits was not explained by improvements in their emotion recognition or affective empathy. [Kimonis and colleagues \(2019\)](#) argue that interventions aiming to remediate the emotional and empathic deficits core to CU traits should be delivered in early childhood when important milestones in emotion recognition and moral development occur: By 3 years old, typically developing children express empathy and caring attitudes ([Dunn et al., 1995](#); [Ellis, 1990](#)) and CU traits can be reliably and validly measured ([Kimonis et al., 2016](#); [Waller et al., 2015b](#); [Wright et al., 2021](#)), with moderate stability into school age (intraclass correlation coefficient [ICC] = 0.93; [Ezpeleta et al., 2013](#)). Emotional literacy training is also more likely to benefit CP + CU children when it targets their specific impairments in recognizing and responding to others’ distress cues (i.e., fear, sad-

ness). For example, directing gaze to the eye region of fearful face stimuli temporarily normalized fear expression recognition in youth with CU traits (Dadds et al., 2008a).

An additional focus of emerging targeted interventions is enhancing warm and responsive parenting, which has been spurred by findings that this parenting style is associated with a reduced risk for elevated CU traits and reduced conduct problems in children with elevated CU traits (Pardini et al., 2007; Pasalich et al., 2011; Waller et al., 2013). Parenting styles promoting greater attachment security (i.e., sensitive responding to child emotion, parental warmth/affection) are well established in the developmental literature as being critical to fostering conscience development and internalization of parental norms, particularly for children with the fearless temperament that is a precursor to CU traits (Barker et al., 2011; Kochanska, 1997). The only other study to our knowledge to exclusively recruit children with a conduct disorder and high CU traits ($N = 40$, 3–8 years) found that parental warmth increased from PMT supplemented with an adjunctive emotional engagement intervention to enhance parent–child reciprocated eye gaze. However, conduct problem and CU trait outcomes were not enhanced relative to PMT supplemented with child-centered play (Dadds et al., 2019). This is likely because PMT programs that integrate a parent–child relationship-building component (i.e., child-centered play) also improve parental warmth and sensitivity (Blizzard et al., 2018).

One PMT program that integrates child-centered play, Parent–Child Interaction Therapy (PCIT; Eyberg & Funderburk, 2011), was adapted to target many of the unique vulnerabilities of children with CU traits (Fleming & Kimonis, 2018). Standard PCIT comprises two distinct, sequential phases delivered via in vivo coaching of parents using a wireless headset from behind a one-way mirror. In standard PCIT, the first child-directed interaction (CDI) phase teaches parents a set of positive parenting skills to improve the parent–child relationship, including use of descriptive praise, speech reflections, behavior imitation and description, and expressions of enjoyment. The second parent-directed interaction (PDI) phase teaches parents to implement a consistent, predictable time-out procedure, used in response to child non-compliance or major rule violations. The callous–unemotional adaptation of PCIT, called PCIT-CU, differs from standard PCIT by (a) systematically coaching parents to use warm and affectionate parenting behaviors with their child during the CDI phase, after providing psychoeducation on its

importance to emotional development for children with CU traits; (b) systematically embedding a dynamic, individualized token economy within the time-out sequence introduced during PDI in order to provide tangible reinforcement of compliant and prosocial behaviors; and (c) targeting the pervasive multilevel emotional processing deficits (Dawel et al., 2012) associated with CU traits using an adjunctive seven-session module delivered after the PDI phase called Coaching and Rewarding Emotional Skills (CARES; see Datyner et al., 2016, for a detailed description). The key treatment objective of CARES is to enhance children’s ability to recognize and prosocially respond to others’ distress via parent modeling, role play, social stories, and consistent reinforcement of children’s emotional skills.

An open trial of PCIT-CU with 23 families with a 3- to 6-year-old child (M age = 4.50 years, $SD = 0.92$) with clinically significant conduct problems and CU traits found that this adapted intervention produced decreases in child conduct problems, CU traits, and increases in empathy, with medium to huge effect sizes ($ds = 0.7–2.0$) that sustained to a 3-month follow-up (Kimonis et al., 2019). By 3 months posttreatment, 75% of treatment completers no longer showed clinically significant conduct problems relative to 25% of dropouts. Treatment retention was high (74%), and parents reported a high level of satisfaction with the program. Although the large improvements in conduct problems produced by the targeted PCIT-CU adaptation appear promising, the open trial design cannot address whether the degree of change was superior to what PCIT could achieve in its standard form, for which a more rigorous randomized controlled trial design is needed.

The Present Study

The primary aim of this proof-of-concept randomized controlled trial was to test the preliminary efficacy and acceptability of PCIT-CU relative to standard PCIT for families with a young child with clinically significant conduct problems and elevated CU traits. We hypothesized that PCIT-CU would be superior to standard PCIT for improving parent-reported child conduct problems and disruptive behavior disorder diagnoses, CU traits, and affective empathy outcomes, as demonstrated by medium effect sizes and indicators of clinically significant and reliable change. We also hypothesized that PCIT-CU and standard PCIT families would be equally compliant with intervention components (i.e., homework completion, attendance), but that parents assigned to PCIT-CU

would have higher treatment retention rates and report greater satisfaction with therapy, supportive of the greater acceptability of the targeted intervention.

Method

PARTICIPANTS

Participants were 43 families with 3- to 7-year-old clinic-referred children (M age = 4.84 years, $SD = 1.12$, 84% boys) with conduct problems and co-occurring CU traits. See Table 1 for a description of sample demographics. See Appendix A for details regarding referral sources and content of targeted study advertisement. Families were eligible to participate if they had a child between 3 and 7 years old who (a) showed elevated CU traits on the Inventory of Callous–Unemotional Traits (ICU) Preschool Version (Kimonis et al., 2016) and (b) scored in the clinically significant range on at least one of the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001) parent-reported externalizing-oriented scales (see the “Eligibility Measures” section). Families were ineligible if (a) the participating caregiver(s) did not speak fluent English, since we were unable to provide standardized clinical and research services and supervision in multiple languages; or (b) the child had received a primary mental health diagnosis other than oppositional defiant disorder (ODD) or conduct disorder (e.g., severe autism spectrum given its overlap with CU traits; Carter Leno et al., 2015); (c) was deaf; or (d) receiving concurrent psychological treatment for behavioral problems as reported by the parent. We permitted children with comorbidities such as attention-deficit/hyperactivity disorder (ADHD) and internalizing disorders to enroll as long as these problems were secondary and less severe than the conduct problems. At the time of the baseline assessment, three children were medicated for ADHD. Study recruitment began in January 2016, and all follow-up assessments were completed by December, 2019. Figure 1 presents the participant flow across the study.

PROCEDURE

The study was approved by the University of New South Wales Human Research Ethics Committee (HC13234) and registered with the Australian New Zealand Clinical Trials Registry (ACTRN12616000280404). We obtained informed consent from all families before they completed a comprehensive assessment taking approximately 2–2.5 hours on five separate occasions: (a) prior to treatment (baseline), (b) follow-

ing the first CDI phase (post-CDI), (c) following the second PDI phase (post-PDI), (d) following the adjunctive CARES module (post-CARES), and (e) 3 months following treatment completion. We conducted follow-up assessments with all families, regardless of dropout—however, despite repeated scheduling attempts, six families failed to complete this assessment, while the remaining three families completed only the online questionnaire component of the assessment. Assessors were not masked to time point. Following the baseline assessment, we randomized participants to receive either standard PCIT ($n = 21$) or PCIT-CU ($n = 22$). Randomization involved the use of sequentially numbered, opaque sealed envelopes and followed an unrestricted simple allocation (Doig & Simpson, 2005). Sample size was contingent on the number of families deemed eligible and that agreed to participate during the recruitment period.

We delivered 1-hour in-person intervention sessions weekly for free to individual families in an urban university research clinic over 21 weeks (PCIT-CU) or 14 weeks (standard PCIT). To standardize treatment dose, standard PCIT participants received weekly telephone consultations for 7 weeks. On average, these calls lasted 14.17 minutes ($SD = 7.12$). Call content included discussion of recent parent or family stressors, behavioral difficulties since the previous call, and PCIT strategies that could be implemented to address behavioral difficulties. The development, refinement, and content of the PCIT-CU protocol is described in detail elsewhere (Fleming & Kimonis, 2018; Kimonis et al., 2019; CARES module: see Datyner et al., 2016; Template for Intervention Description and Replication [TIDieR] checklist provided in Appendix A). All families participated in one parent-only teach session and six parent–child coach sessions each for the CDI, PDI, and CARES (PCIT-CU group only) phases. We selected this fixed format approach over the typical variable treatment length format of standard PCIT because it yielded improved outcomes and lower attrition rates for complex high-risk populations (Thomas & Zimmer-Gembeck, 2012). The average length of treatment for treatment completers was 39.12 weeks ($SD = 9.35$) for standard PCIT and 35.40 weeks ($SD = 7.85$) for PCIT-CU. One constant caregiver, typically the mother, completed all treatment/assessment sessions (Appendix A provides details of father involvement). Therapists ($n = 7$, 86% woman identifying) were licensed, clinically trained psychologists who received intensive in vivo training from the last/senior author, who is the PCIT-CU developer and a certified

Table 1
Demographic Characteristics of Participants for the Overall Sample and by Group

Variable	Overall sample <i>M (SD)</i>	Standard PCIT <i>M (SD)</i>	PCIT-CU <i>M (SD)</i>	Significance test
Child age, years	<i>N</i> = 43 4.84 (1.12)	<i>n</i> = 21 4.49 (0.99)	<i>n</i> = 22 5.16 (1.15)	<i>t</i> (41) = -2.04, <i>p</i> = .048*, <i>d</i> = 0.62
Maternal age, years	<i>N</i> = 43 38.30 (5.35)	<i>n</i> = 21 38.76 (3.43)	<i>n</i> = 22 37.86 (6.76)	<i>t</i> (41) = 0.55, <i>p</i> = .59
Paternal age, years	<i>N</i> = 41 42.56 (6.38)	<i>n</i> = 21 42.86 (5.69)	<i>n</i> = 20 42.25 (7.17)	<i>t</i> (39) = 0.30, <i>p</i> = .77
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	
Child sex	<i>N</i> = 43	<i>n</i> = 21	<i>n</i> = 22	Fisher's exact test: <i>p</i> = 1.00
Male	36 (83.7)	18 (85.7)	18 (81.8)	
Female	7 (16.3)	3 (14.3)	4 (18.2)	
Maternal race/ethnicity	<i>N</i> = 43	<i>n</i> = 21	<i>n</i> = 22	$\chi^2(1) = 0.11$, <i>p</i> = .74
Caucasian	29 (67.4)	14 (66.7)	15 (68.2)	
Asian	8 (18.6)	3 (14.3)	5 (22.7)	
Middle Eastern	1 (2.3)	1 (4.8)	0 (0.0)	
African	1 (2.3)	0 (0.0)	1 (4.5)	
Other race/ethnicity	3 (7.0)	3 (14.3)	0 (0.0)	
Missing	1 (2.3)	0 (0.0)	1 (4.5)	
Paternal race/ethnicity	<i>N</i> = 42	<i>n</i> = 21	<i>n</i> = 21	$\chi^2(1) = 53$, <i>p</i> = .47
Caucasian	32 (74.4)	15 (71.4)	17 (77.3)	
Asian	5 (11.6)	2 (9.5)	3 (13.6)	
Pacific Islander	1 (2.3)	0 (0.0)	1 (4.5)	
Middle Eastern	1 (2.3)	1 (4.8)	0 (0.0)	
Other race/ethnicity	3 (7.0)	3 (14.3)	0 (0.0)	
Missing ^a	1 (2.3)	0 (0.0)	1 (4.5)	
Parent marital status	<i>N</i> = 43	<i>n</i> = 21	<i>n</i> = 22	Fisher's exact test: <i>p</i> = .61
Never married	1 (2.3)	1 (4.8)	0 (0.0)	
Married	32 (74.4)	18 (85.7)	14 (63.6)	
Cohabiting	6 (14.0)	2 (9.5)	4 (18.2)	
Divorced/separated	4 (9.3)	0 (0.0)	4 (18.1)	
Annual household income	<i>N</i> = 41	<i>n</i> = 20	<i>n</i> = 21	$\chi^2(1) = 1.21$, <i>p</i> = .27
≤\$160,000	21 (48.8)	12 (57.1)	9 (40.9)	
>\$160,000	20 (46.5)	8 (38.1)	12 (54.5)	
Missing ^b	2 (4.7)	1 (4.8)	1 (4.5)	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	
	<i>N</i> = 43	<i>n</i> = 21	<i>n</i> = 22	
ECBI				
Intensity	178.60 (22.89)	180.95 (25.06)	176.36 (20.96)	<i>t</i> (41) = 0.65, <i>p</i> = .52
Problem	25.14 (5.91)	25.76 (5.78)	24.55 (6.12)	<i>t</i> (41) = 0.67, <i>p</i> = .51
CBCL <i>T</i> scores				
Aggressive Behavior	77.30 (10.79)	79.71 (11.57)	75.00 (9.70)	<i>t</i> (41) = 1.45, <i>p</i> = .15
Externalizing	72.95 (8.60)	74.57 (9.62)	71.41 (7.41)	<i>t</i> (41) = 1.21, <i>p</i> = .23
Oppositional Defiant	72.60 (6.85)	73.00 (8.06)	72.23 (5.64)	<i>U</i> = 189, <i>z</i> = -1.04, <i>p</i> = .30
Attention-Deficit/Hyperactivity	63.70 (8.04)	63.14 (8.60)	64.23 (7.63)	<i>t</i> (41) = -0.44, <i>p</i> = .66
ICU total	36.14 (9.82)	39.14 (9.80)	33.27 (9.14)	<i>t</i> (41) = 2.03, <i>p</i> = .049*, <i>d</i> = 0.62
GEM Affective	-6.02 (13.58)	-11.95 (12.92)	-0.36 (11.89)	<i>t</i> (41) = -3.06, <i>p</i> = .004*, <i>d</i> = 0.93

Note. PCIT = parent-child interaction therapy; CU = callous-unemotional; *M* = mean; *SD* = standard deviation; *d* = Cohen's *d*; ECBI = Eyberg Child Behavior Inventory; CBCL = Child Behavior Checklist; ICU = Inventory of Callous-Unemotional Traits; GEM = Griffith Empathy Measure.

^a One family did not report employment status.

^b Two families did not report income data. Income data are reported in Australian dollars.

* *p* < .05.

PCIT trainer, involving co-treatment roles on two cases and regular clinical supervision to maintain

treatment fidelity. Therapists administered intervention to 6.1 cases on average (*SD* = 5.9, range:

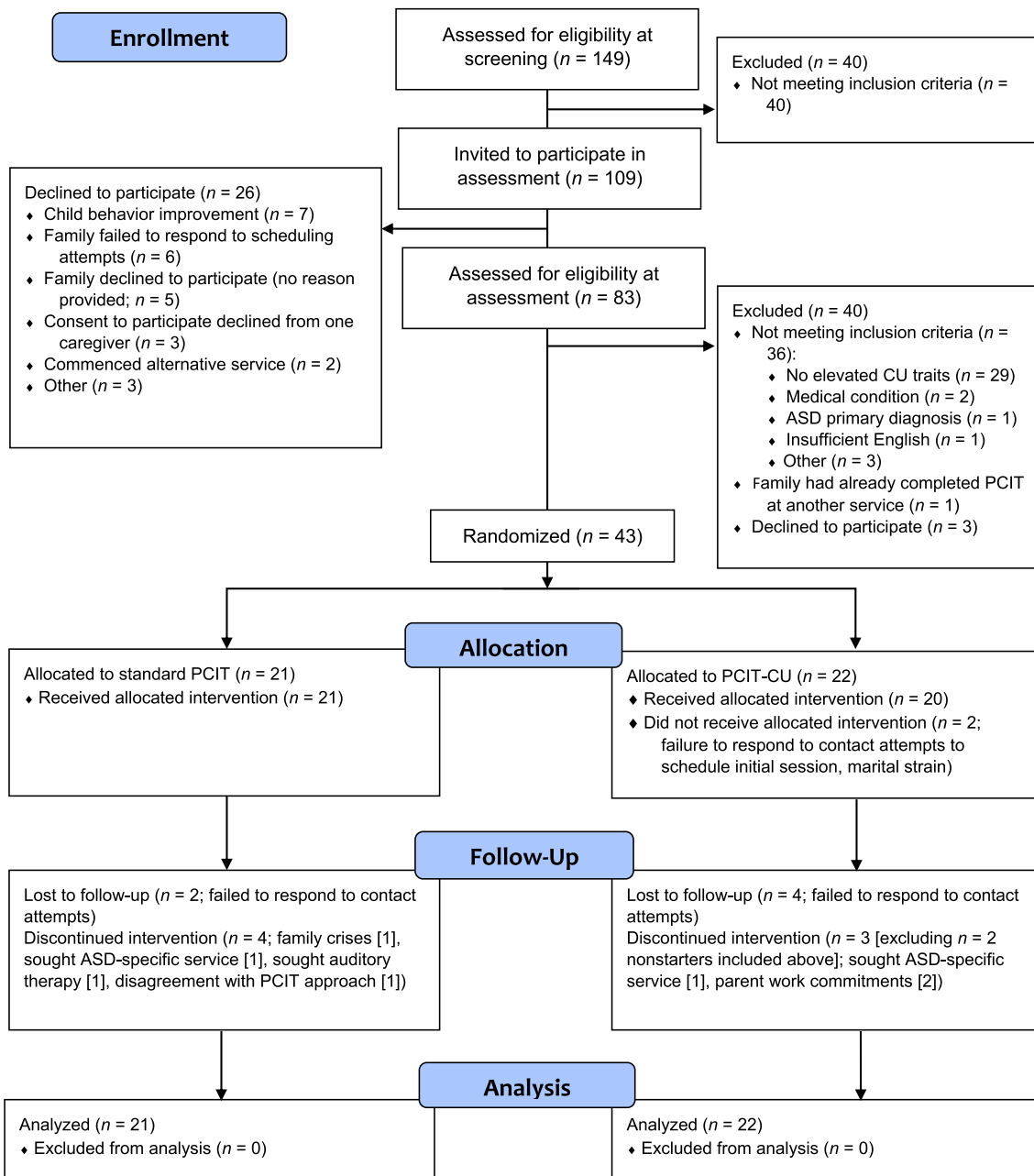


FIGURE 1 CONSORT flow of participants. Note. CU = callous–unemotional; ASD = autism spectrum disorder; PCIT = parent–child interaction therapy.

1–18) and delivered treatment across both conditions.

MEASURES

Eligibility Measures

One caregiver completed brief measures of CU traits and conduct problems during the initial intake telephone call to determine eligibility. Since it was important for this call to be brief, we used 10 items from the ICU (see the “Outcome Measures” section) to screen for CU traits (e.g., “Feels

bad or guilty when he/she has done something wrong” [reverse scored]). We selected these items because they reflect the four criteria of the DSM-5 LPE specifier. Parents rated children on a 4-point scale (0 = *not at all true*, 1 = *somewhat true*, 2 = *very true*, 3 = *definitely true*). We considered CU traits as “present” when at least two of four LPE criteria were endorsed, as indicated by a 2 or 3 rating on its respective ICU item(s), which was determined to be the optimum way to assess symptom presence by [Kimonis et al. \(2015\)](#). In

the current study, this method identified a sample with average baseline 24-item total ICU scores (M score = 36.14, SD = 9.82) that fell 1.5 standard deviation units above the mean for a large sample (N = 622) of preschoolers (Ezpeleta et al., 2013). We considered conduct problems to be “clinical” if at least one of the externalizing-oriented scales on the appropriate age version of the ASEBA Child Behavior Checklist (CBCL) was in the clinical range (T scores ≥ 70 for syndrome/DSM-oriented scales and ≥ 64 for composite scale).

Outcome Measures

Participants completed the following measures at each of the five assessment points. Appendix A provides details regarding the administration, scoring, and established psychometric properties of each measure. For all outcome measures, we combined mother and father scores in a conservative fashion by taking the higher rating between raters (i.e., “resolved” score) to circumvent potential underreporting of problems at baseline and overreporting of treatment effect at posttreatment. We calculated the ICC to determine interrater agreement between caregivers. We used the two-way random average measures form for ICC and selected absolute agreement when calculating the ICC. ICCs below 0.40 were considered poor, 0.40–0.59 were fair, 0.60–0.74 good, and 0.75 or above excellent (Cicchetti, 1994).

Conduct Problems. We assessed child conduct problems using the Intensity and Problem scales of the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999). In the current study, both Intensity and Problem scores showed good to excellent internal consistency (Cronbach’s α s = .88–.95 and .80–.92, respectively). The ICC between mother and father ECBI Intensity and Problem scale scores ranged between 0.49–0.87 and 0.43–0.75, respectively, across all time points, with the exception of the Problem scale at the post-PDI assessment, which demonstrated poor agreement.

We also assessed conduct problems using resolved T scores from selected syndrome and DSM-oriented scales of the ASEBA CBCL/1.5–5 and CBCL/6–18 (Achenbach & Rescorla, 2001), including the Aggressive Behavior, DSM-Oriented Oppositional Defiant Problems, and composite Externalizing scales. In the current study, resolved scale scores demonstrated acceptable to excellent internal consistency across time (Cronbach’s α s for CBCL1.5–5 scales: Aggressive Behavior = .87–.96, Oppositional Defiant Problems = .76–.93, and Externalizing = .86–.95; α s for CBCL/6–18 scales Aggressive Behavior = .77–

.87; Externalizing = .83–.87), except the CBCL/6–18 Oppositional Defiant Problems scale which was low (Cronbach’s α s = .33–.62). ICCs between mother and father CBCL Aggressive Behavior, Oppositional Defiant Problems, and Externalizing T scores ranged between 0.22–0.54, 0.31–0.67, and 0.38–0.63, respectively, across time.

Diagnostic Status. We assessed child ODD and CD diagnostic status using the ODD and CD modules from the semistructured Diagnostic Interview Schedule for Children, Adolescents, and Parents (DISCAP; Holland & Dadds, 1997), conducted with the primary caregiver to assess for the presence of DSM childhood mental disorders. We used the categorical (diagnosis/no diagnosis) data in the current study.

Callous–Unemotional Traits. We assessed CU traits using 24-item total scores from the preschool version of the ICU. In the current study, total resolved scores showed good to excellent internal consistency across all time points (Cronbach’s α s = .85–.91). The ICC between mother and father 24-item ICU total scores ranged between 0.44 and 0.78 across all time points.

Affective Empathy. We assessed affective empathy using the Affective Empathy subscale from the Griffith Empathy Measure (GEM; Dadds et al., 2008b). In the current study, resolved GEM Affective Empathy scale scores demonstrated good to excellent reliability (Cronbach’s α s = .84–.90). ICCs between mother and father GEM Affective Empathy scores ranged between 0.56 and 0.74.

Treatment Acceptability. We assessed treatment acceptability using the Therapy Attitude Inventory (TAI; Brestan et al., 1999). In the current study, total TAI scores demonstrated good internal consistency at post-CARES (α = .89) and follow-up (α = .87). ICCs between mother and father TAI total scores were 0.56 at post-CARES and 0.44 at follow-up. Indicators of treatment engagement included premature attrition from treatment, session attendance, and weekly homework compliance.

PLANNED ANALYSES

See Appendix A for planned analyses for missing data and baseline group differences.

Efficacy

We conducted linear mixed-effects modeling in SAS Version 9.4 using data from the full intent-to-treat sample (N = 43) to examine the efficacy of PCIT-CU relative to standard PCIT over time for continuous outcome variables, using restricted

maximum likelihood estimation to account for missing data. For each outcome variable, we estimated a series of latent growth curve models: We included the fixed effects of time (weeks since baseline), treatment condition, and the interaction of Time \times Condition were included as predictors and explored the shape of change over time by examining linear, quadratic, and cubic change. We examined three random effects covariance structures for each model, including random intercept, random intercept within each group, and random intercept and separate error variances for each assessment time point. We examined two covariance matrices: Compound symmetry and autoregressive(1). For each outcome variable, we retained the model with the lowest model fit values (AIC and BIC). We included child age, baseline ICU 24-item total scores (except the model predicting ICU 24-item total scores), baseline GEM affective scores (except the model predicting GEM affective scores), and therapist as covariates in all models. For each model, we specified planned group differences in change over time for several time intervals (see Table 2 and Supplemental Tables 2–8) and calculated effect sizes (growth model analysis [GMA] d) for each time interval. We conducted binomial logistic regression to evaluate group differences diagnostic status at post-CARES and follow-up. Assumptions were checked and satisfied for all linear mixed-effects and binomial logistic regression models.

Clinical Significance of Efficacy Outcomes

To examine clinically significant differences between treatment conditions, we used descriptive analyses to identify the proportion of treatment completers in each treatment group who demonstrated Reliable and clinically significant change. We calculated Reliable Change Index (RCI) scores to determine whether the magnitude of individual-level change on parent-rated measures from pretreatment to posttreatment exceeded the margin of measurement error: (pretest scale score – posttest scale score)/ SE_{diff} . Individual RCI scores > 1.96 (reliable improvement) were considered statistically significant, $p < .05$ (Jacobson & Truax, 1991). We determined clinically significant change by identifying the treatment completers who no longer showed clinically significant conduct problems by post-CARES and follow-up based on ECBI and CBCL T scores and rates of DISCAP ODD/CD diagnoses.

Treatment Acceptability

To examine group differences in treatment acceptability, we conducted binomial logistic regression to evaluate group differences in treatment attri-

tion, and multiple linear regression to evaluate group differences in treatment satisfaction, total number of missed or canceled treatment sessions, and average rates of homework completion as reported by treatment completers. Assumptions were checked and satisfied for all binomial logistic and multiple linear regression models, with the exception of identification of a single outlier from the PCIT-CU group in the regression model predicting cancellations. Results regarding group differences remained unchanged when we removed the outlier.

Results

MISSING DATA

There were no significant associations between missingness and demographic variables or baseline ECBI Intensity scores at any time point for the ECBI, CBCL, ICU, or GEM, except for a significant association between missingness and child age at the post-CARES assessment time point for the ECBI, CBCL, and GEM. This association disappeared when we removed outliers.

BASELINE DIFFERENCES

Table 1 presents descriptive information and analyses for demographic characteristics of participants for the full sample and by treatment condition. Regarding demographic variables, there were no significant differences between standard PCIT and PCIT-CU groups in parent age, race/ethnicity (binarized: non-Caucasian/Caucasian), marital status (binarized: currently in relationship/not in relationship), household income (binarized using median income), or child sex. There was a significant difference between standard PCIT and PCIT-CU groups in child age ($p = .048$), such that children in the PCIT-CU condition were significantly older than children in standard PCIT. Regarding treatment outcome variables, there were no significant differences between standard PCIT and PCIT-CU groups in baseline ECBI Intensity scores or CBCL Aggressive Behavior, Externalizing, or DSM Oppositional Defiant Problems T scores. There were significant differences between conditions in baseline ICU ($p = .049$) and GEM Affective ($p = .004$) scores, such that children in standard PCIT were significantly higher in CU traits and significantly lower in empathy than children in PCIT-CU. Regarding baseline diagnostic status, there were no significant group differences in the proportion of children diagnosed with ODD or CD on the DISCAP: In the standard PCIT group, 10 children were diagnosed with ODD only and 8 children with ODD and CD. In

Table 2
Results of Mixed-Effects Models Examining the Linear and Quadratic Effects of Time, Group, and Their Interactions on Outcomes

Estimates and <i>p</i> values										
Variable	Time		Time ²		Group		Time × Group		Time ² × Group	
	<i>b</i> [95% CI]	<i>p</i>	<i>b</i> [95% CI]	<i>p</i>	<i>b</i> [95% CI]	<i>p</i>	<i>b</i> [95% CI]	<i>p</i>	<i>b</i> [95% CI]	<i>p</i>
ECBI Intensity	−4.68 [−5.70, −3.65]	<.001*	0.10 [0.07, 0.13]	<.001*	6.95 [−18.38, 32.29]	.59	0.23 [−1.21, 1.66]	.75	−0.01 [−0.05, 0.02]	.49
ECBI Problem	−1.14 [−1.40, −0.89]	<.001*	0.02 [0.02, 0.03]	<.001*	−1.23 [−7.00, 4.54]	.67	0.38 [0.03, 0.73]	.03*	−0.01 [−0.02, −0.00]	.01*
CBCL Agg	−1.39 [−1.77, −1.00]	<.001*	0.03 [0.02, 0.04]	<.001*	−6.61 [−14.58, 1.36]	.10	0.51 [−0.03, 1.06]	.06**	−0.02 [−0.03, 0.00]	.05**
CBCL Ext	−1.11 [−1.46, −0.76]	<.001*	0.02 [0.01, 0.03]	<.001*	−2.26 [−9.00, 4.48]	.50	0.45 [−0.04, 0.93]	.07**	−0.01 [−0.03, −0.00]	.05**
CBCL Opp Def	−1.02 [−1.36, −0.68]	<.001*	0.02 [0.01, 0.03]	<.001*	−0.40 [−6.16, 5.36]	.89	0.22 [−0.25, 0.68]	.36	−0.01 [−0.02, 0.01]	.38
ICU 24-item total	−0.55 [−0.84, −0.26]	<.001*	0.01 [0.00, 0.02]	.004*	−2.20 [−10.38, 5.98]	.60	0.31 [−0.10, 0.72]	.13	−0.01 [−0.02, 0.00]	.09**
GEM Affective	0.44 [0.10, 0.78]	.01*	−0.01 [−0.02, −0.00]	.01*	11.19 [2.52, 19.86]	.01*	−0.45 [−0.93, 0.02]	.06**	0.01 [−0.00, 0.02]	.08**

Group differences in changes in scores for specific time intervals

Variable	Entire study period (Baseline to 3MFU)			Active treatment period (Baseline to post-CARES)			Follow-up period (Post-CARES to 3MFU)		
	<i>b</i> [95% CI]	<i>p</i>	<i>d</i>	<i>b</i> [95% CI]	<i>p</i>	<i>d</i>	<i>b</i> [95% CI]	<i>p</i>	<i>d</i>
ECBI Intensity	−9.14 [−27.33, 9.05]	.32	−0.40	−2.23 [−18.18, 13.71]	.78	−0.10	−6.90 [−19.65, 5.84]	.29	−0.30
ECBI Problem	−2.35 [−7.20, 2.49]	.34	−0.40	2.02 [−1.83, 5.87]	.30	0.34	−4.37 [−7.88, −0.85]	.02*	−0.73
CBCL Agg	−0.96 [−7.30, 5.38]	.77	−0.09	3.69 [−1.94, 9.32]	.20	0.35	−4.65 [−9.72, 0.42]	.07**	−0.44
CBCL Ext	−1.07 [−6.43, 4.29]	.69	−0.13	3.10 [−1.92, 8.13]	.22	0.36	−4.17 [−8.46, 0.11]	.06**	−0.49
CBCL Opp Def	0.51 [−4.66, 5.67]	.85	0.07	1.96 [−2.87, 6.79]	.42	0.28	−1.45 [−5.60, 2.69]	.49	−0.21
ICU 24-item total	−1.21 [−5.78, 3.35]	.60	−0.13	1.96 [−2.30, 6.23]	.36	0.21	−3.18 [−6.80, 0.45]	.09**	−0.34
GEM Affective	−1.59 [−6.78, 3.60]	.55	−0.13	−4.33 [−9.27, 0.60]	.08**	−0.35	2.74 [−1.37, 6.86]	.19	0.22

Note. Time = weeks since baseline; Group = allocation to treatment condition; ECBI = Eyberg Child Behavior Inventory; CBCL = Child Behavior Checklist; Agg = aggressive behavior; Ext = externalizing; Opp Def = oppositional defiant problems; ICU = Inventory of Callous–Unemotional Traits; GEM = Griffith Empathy Measure; 3MFU = 3-month follow-up assessment; CARES = Coaching and Rewarding Emotional Skills; *d* = growth model analysis (GMA) *d* effect size. See Supplemental Materials for bootstrapped confidence intervals for each GMA *d*, which were obtained from 2,000 bootstrapped samples, calculating the effect size and using the 2.5th and 97.5th percentile values (Feingold, 2019). Models control for child age in years, baseline ICU 24-item total scores, baseline GEM Affective scores, and therapist.

* *p* < .05.

** *p* < .10.

the PCIT-CU group, 12 children were diagnosed with ODD only, 1 child with CD only, and 7 children with ODD and CD.

EFFICACY

Results of linear mixed-effects models are presented in Table 2, with additional details provided in Supplemental Tables 2–8. Plots of model-predicted means are presented in Figure 2.

Conduct Problems

Regarding group differences in changes in conduct problem scores over time, condition did not interact significantly with the linear or quadratic effects

of time for ECBI Intensity or CBCL Oppositional Defiant Problems scores (Table 2). However, there were significant group differences in linear and quadratic change over time for ECBI Problem scores, indicating that the shape of change over time for the standard PCIT group differed from that of the PCIT-CU group for this outcome, such that scores appeared to deteriorate for the standard PCIT group (Figure 2). CBCL Externalizing and Aggressive Behavior scores demonstrated a similar pattern at a trend level (both $p = .05$; Table 2). Planned comparisons of the group differences in mean change during the follow-up period

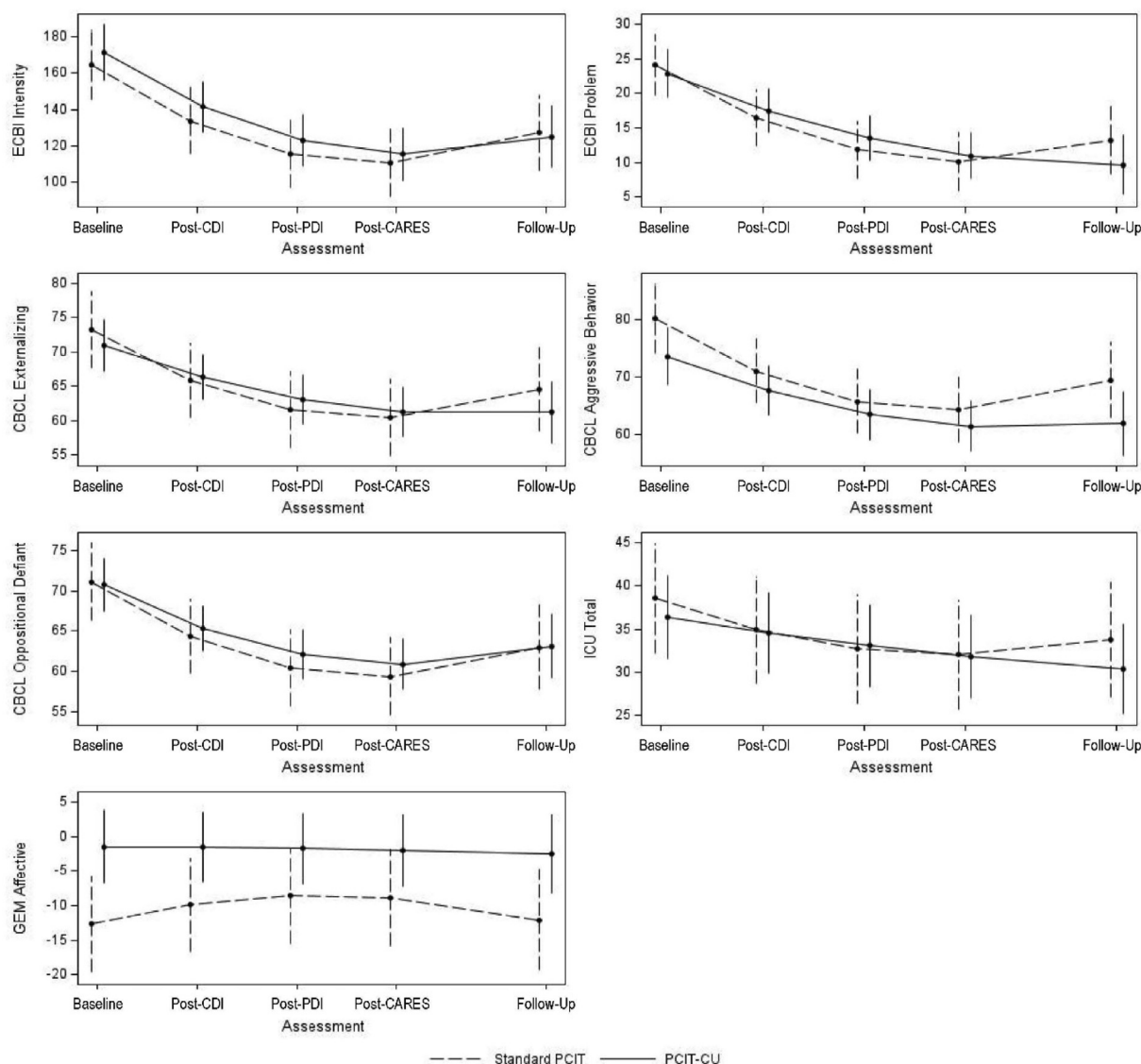


FIGURE 2 Plots of model-predicted means. Note. ECBI = Eyberg Child Behavior Inventory; CBCL = Child Behavior Checklist; ICU = Inventory of Callous–Unemotional Traits; GEM = Griffith Empathy Measure; CDI = child-directed interaction; PDI = parent-directed interaction; CARES = coaching and rewarding emotional skills. Plots of model-predicted means shown for primary outcomes at each assessment time point for standard parent–child interaction therapy (PCIT; dashed line) and PCIT callous–unemotional (PCIT-CU; solid line).

were statistically significant for ECBI Problem scores and trending toward significance for CBCL Aggressive Behavior ($p = .07$) and Externalizing ($p = .06$) scores (Table 2). Children in the PCIT-CU group were rated as maintaining improvement or continuing to improve during the follow-up period, while children in the standard PCIT group deteriorated (Figure 2). The magnitude of these between-group effects were moderate for ECBI Problem scores ($d = -0.73$) and small to moderate for CBCL Aggressive Behavior ($d = -0.44$) and Externalizing ($d = -0.49$; Table 2). Regarding the other time intervals of interest, no other significant group differences in rates of change in conduct problems over time were found. See Appendix A for interpretation of nonsignificant trends toward group differences in rate of change during these other time intervals.

Diagnostic Status

Supplemental Table 9 presents logistic regression results of condition on diagnostic status at post-CARES and follow-up, controlling for covariates. The models predicting diagnostic status were not significant at post-CARES, $\chi^2(4) = 1.90$, $p = .75$, or follow-up, $\chi^2(4) = 1.76$, $p = .78$. Condition did not predict diagnostic status at either time point.

Callous–Unemotional Traits

Regarding child CU traits, group did not interact significantly with the linear or quadratic effects of time, although there was a nonsignificant trend toward a group difference in quadratic change over time ($p = .09$; Table 2). Planned comparisons indicated a nonsignificant trend toward group differences in rate of change during the follow-up period for ICU scores ($p = .09$), such that children in PCIT-CU continued to improve after treatment while children in standard PCIT deteriorated during this period (Figure 2), with a small effect size ($d = -0.34$; Table 2).

Affective Empathy

Regarding child affective empathy, condition did not interact significantly with the linear or quadratic effect of time, although there were nonsignificant trends toward group differences in the linear ($p = .06$) and quadratic ($p = .08$) effects of time (Table 2). Planned comparisons indicated nonsignificant trends toward group differences in rate of change in GEM Affective Empathy scores during several time intervals, including the active treatment period ($p = .08$). Children in standard PCIT were rated as improving during these time intervals, while children in PCIT-CU did not improve during these intervals (Figure 2). All of

these nonsignificant effects were very small to small in magnitude ($ds = -0.12$ to -0.35 ; Table 2).

CLINICAL SIGNIFICANCE OF EFFICACY OUTCOMES

Table 3 presents percentages of participants who showed reliable and clinically significant change in outcomes at post-CARES and follow-up. Regarding group differences in outcomes measuring conduct problems, the majority of PCIT-CU participants (88%) were rated as reliably improved in the frequency and number of conduct problems at both time points, while percentages were more varied for the standard PCIT group in frequency and number of conduct problems at post-CARES (87% and 73%, respectively) and follow-up (71% and 53%, respectively). Reliable improvement in conduct problems measured using the CBCL Externalizing scales was more equivocal (Table 3). Half or more of all participants were rated as below the clinical cutoff on all measures of conduct problems at post-CARES and follow-up, with the exception of ECBI Problem scores for the standard PCIT group at follow-up. Rates of reliable improvement were considerably lower for parent-rated affective outcomes than conduct problem outcomes. Specifically, less than a third of participants in standard PCIT were rated as showing reliable improvement in child CU traits and affective empathy at post-CARES and follow-up, with similarly low rates of reliable improvement found for the PCIT-CU group at both time points.

TREATMENT ACCEPTABILITY

Results of analyses are described below, with full details provided in Supplemental Tables 9 and 10. Treatment retention was high in both standard PCIT and PCIT-CU groups, with four (19%) families in the standard PCIT group and five (23%) families in the PCIT-CU group prematurely dropping out of treatment. See Figure 1 for details regarding attrition. Supplemental Table 9 presents logistic regression results of group on treatment attrition, controlling for covariates. The model predicting treatment attrition was not significant, $\chi^2(4) = 3.33$, $p = .51$. Condition did not predict treatment attrition.

Regarding treatment satisfaction reported by treatment completers, raw mean (SD) TAI scores for standard PCIT and PCIT-CU conditions, respectively, were 4.48 ($SD = 0.41$) and 4.49 ($SD = 0.47$) at post-CARES and 4.42 ($SD = 0.36$) and 4.53 ($SD = 0.42$) at 3-month follow-up, which correspond to a high level of satisfaction with the

Table 3
Number (Percentage) of Treatment-Completing Participants With Reliable and Clinically Significant Improvement at Posttreatment and 3-Month Follow-Up

Variable	Overall <i>N</i> = 34		Standard PCIT <i>n</i> = 17		PCIT-CU <i>n</i> = 17	
	Post	3MFU	Post	3MFU	Post	3MFU
	Reliable change					
ECBI Intensity	27 (87)	26 (79)	13 (87)	12 (71)	14 (88)	14 (88)
ECBI Problem	25 (81)	23 (70)	11 (73)	9 (53)	14 (88)	14 (88)
CBCL Agg	19 (61)	18 (60)	8 (53)	8 (50)	9 (56)	9 (64)
CBCL Ext	19 (61)	16 (53)	10 (67)	7 (44)	9 (56)	9 (64)
CBCL Opp Def	18 (58)	16 (53)	8 (53)	9 (56)	10 (63)	7 (50)
ICU 24-item total	10 (33)	8 (28)	4 (27)	4 (27)	6 (40)	4 (29)
GEM Affective	3 (10)	4 (13)	2 (13)	3 (19)	1 (6)	1 (7)
	Clinically significant change					
ECBI Intensity	18 (58)	17 (52)	9 (60)	9 (53)	9 (56)	8 (50)
ECBI Problem	22 (71)	16 (49)	11 (73)	7 (41)	11 (69)	9 (56)
CBCL Agg	25 (81)	19 (63)	12 (80)	9 (56)	13 (81)	10 (71)
CBCL Ext	18 (58)	15 (50)	8 (53)	8 (50)	10 (63)	7 (50)
CBCL Opp Def	25 (81)	20 (67)	11 (73)	11 (69)	14 (88)	9 (64)
DISCAP ODD	17 (61)	20 (69)	9 (64)	11 (69)	8 (57)	9 (69)
DISCAP CD	23 (89)	27 (93)	12 (86)	16 (100)	11 (92)	11 (85)
DISCAP ODD/CD	16 (57)	19 (66)	8 (57)	11 (69)	8 (57)	8 (62)

Note. PCIT = parent-child interaction therapy; CU = callous-unemotional; Post = post-CARES (Coaching and Rewarding Emotional Skills); 3MFU = 3-month follow-up; ECBI = Eyberg Child Behavior Inventory; CBCL = Child Behavior Checklist; Agg = aggressive behavior; Ext = externalizing; Opp Def = oppositional defiant problems; ICU = Inventory of Callous-Unemotional Traits; GEM = Griffith Empathy Measure; DISCAP = Diagnostic Interview Schedule for Children, Adolescents, and Parents; ODD/CD = oppositional defiant disorder/conduct disorder.

process and outcome of therapy. Supplemental Table 10 presents separate multiple linear regression results of condition on TAI scores at post-CARES and follow-up, controlling for covariates. The model predicting TAI scores was not significant at post-CARES, $F(4, 19) = 0.28$, $p = .86$, $R^2 = 0.06$, or follow-up, $F(4, 22) = 0.61$, $p = .66$, $R^2 = 0.10$. Treatment condition did not predict TAI scores at post-CARES or follow-up.

Regarding other indicators of engagement for treatment completers, raw mean (*SD*, range) number of missed or canceled sessions for standard PCIT and PCIT-CU groups were 2.65 (2.60, 0–9) and 4.35 (3.97, 0–14), respectively, while raw mean (*SD*) rates of total homework completion were 64% (16.5) and 61% (13.8), respectively. Supplemental Table 10 presents separate multiple linear regression results of group on total number of session cancellations and rates of homework completion, controlling for covariates. The model predicting total number of session cancellations was not significant, $F(4, 29) = 1.30$, $p = .29$, $R^2 = 0.15$, nor was the model predicting homework completion, $F(4, 29) = 0.55$, $p = .70$, $R^2 = 0.07$. Condition did not predict session cancellations or homework completion.

Discussion

The purpose of this research was to provide an initial randomized controlled trial of an evidenced-based parent training intervention adapted to target the unique developmental needs of young CP + CU children and to test whether this led to improved outcomes relative to the standard version of parent training. Findings indicated that both standard PCIT and its adaptation for CU traits produced significant improvements in child conduct problem symptoms and CU traits. In partial support of hypotheses, children assigned to PCIT-CU showed sustained improvement in conduct problems to 3-month follow-up, whereas standard PCIT participants showed some deterioration in outcomes. We found a similar pattern of group differences during the follow-up period for child CU traits, although this effect did not reach statistical significance. Contrary to hypotheses, there were no group differences in treatment acceptability, which was high for both conditions. We discuss four key study findings and their implications in greater depth below.

Our first key finding was that all children showed significant improvements in parent-rated and diagnosed conduct problems, regardless of

treatment condition. While this finding was contrary to hypotheses, it corroborates the field's shifting perspective on the relationship between child CU traits and response to parenting interventions: The deleterious effect of elevated baseline CU traits on treatment efficacy is not uniform. Specifically, studies failed to identify a negative effect of CU traits on treatment outcomes with very young or subclinical samples (e.g., Hyde et al., 2013) and when interventions were intensive (>20 sessions) and addressed multiple risk factors (e.g., Kolko & Pardini, 2010). Together, these studies suggested that the detrimental impact of CU traits on treatment outcomes may be ameliorated by early, intensive intervention. It remains unclear where PCIT falls along this spectrum of intervention intensity. While some evidence supports that PCIT becomes less efficacious and acceptable as child CU traits increase (Fleming et al., 2020; Kimonis et al., 2014), current findings suggest that standard PCIT is sufficiently intense to overcome the impact of CU traits on treatment outcomes during the active treatment period. One possible explanation for PCIT's efficacy is its parent-child relationship-building component—that is, the standard version of PCIT's CDI phase may sufficiently ameliorate low parental warmth and responsiveness (Blizzard et al., 2018), which is a risk factor for CU traits (Waller et al., 2013).

On the other hand, whereas families that received standard PCIT experienced considerable intervention benefit commensurate in magnitude with the PCIT-CU condition during the active treatment period, they showed deterioration in outcomes once treatment ceased. In contrast, children in the PCIT-CU group maintained their gains or continued to improve during the follow-up period. This differential pattern of maintenance occurred across parent-rated outcomes, including number of conduct problems and frequency of severe conduct problems, although some of these effects were trend level ($ps = .05-.07$), which likely occurred due to small sample size resulting in low power. Thus, our second key finding is that standard PCIT may produce short-term improvements that then fade out, suggesting it inadequately targets mechanisms responsible for sustained treatment gains.

Investigating the mechanisms by which PCIT-CU might enhance maintenance of treatment gains is an important next step in efforts to validate PCIT-CU as an evidence-based intervention for young CP + CU children. PCIT-CU is assumed to work because it improves parental warmth, engages reward-dominant learning, and/or

improves child emotional skills. In preliminary support of this hypothesis, current results indicated that PCIT-CU might sustain improvement in severe conduct problems better than standard PCIT. Specifically, the two treatments produced similar improvements in mild conduct problems (e.g., emotional dysregulation and noncompliance), as assessed by the ECBI Intensity and CBCL Oppositional Defiant Problems scales, but PCIT-CU was more effective than standard PCIT in maintaining improvement in severe conduct problems (e.g., aggression and stealing), as assessed by the CBCL Aggressive Behavior and Externalizing scales.

Mechanistically, standard PCIT's discipline procedure reduces emotional dysregulation and noncompliance by extinguishing these attention- and escape-maintained behaviors via time-out. Severe behaviors, such as physical aggression, may be less amenable to time-out because they receive immediate reinforcement, according to the differential amygdala activation model of psychopathy (Moul et al., 2012). This model predicts that behaviors (e.g., pushing another child to get a desired toy) are reinforced as long as they produce a positive outcome for the perpetrator (e.g., getting the toy), irrespective of later punishment or the negative impact of the behavior on others (e.g., injury), because the child focuses on the general affective valence of an outcome (e.g., good, reward) rather than its specific properties (e.g., "I feel good *because* I have the toy"). This tendency prevents chaining of associations that are critical for the individual to learn from consequences (e.g., "Max is now sad because I have the toy"). Thus, standard PCIT's typical discipline sequence may be less effective for producing lasting change in severe conduct problems among CP + CU children. Unlike standard PCIT, PCIT-CU tailors to the reward-driven and punishment-insensitive response styles of CP + CU children (Barker et al., 2011; Byrd et al., 2014) by combining time-out with tangible rewards for compliance and prosocial behaviors, and explicitly teaching and reinforcing application of emotional skills. These strategies may be more effective than time-out alone in sustaining CP + CU children's motivation to inhibit aggressive, destructive, and right-violating behaviors.

However, it is also possible that PCIT-CU works via other, as-yet-unidentified mechanistic pathways. In preliminary support of this hypothesis, PCIT-CU was more effective than standard PCIT in sustaining improvement in parent *perceptions* of child conduct problems. Specifically, standard PCIT was associated with deterioration in

ECBI Problem scale score gains, while PCIT-CU showed maintenance. Since this scale represents parents' tolerance of disruptive behaviors (Eyberg & Pincus, 1999), this finding suggests that PCIT-CU may more effectively shift parental cognitions or attributions about the stability, malleability, or intentionality of children's conduct problems. Critically, these dimensions are known to elicit negative affective and behavioral parent responses to disruptive child behavior (Morrissey-Kane & Prinz, 1999). Since research supports the importance of parental attributions as both a target and moderator of treatment response (Sawrikar & Dadds, 2018)—perhaps especially among CP + CU children (Palm et al., 2019)—future investigation of this potential mechanism of sustained gains following PCIT-CU is warranted.

Our third key finding concerns affective treatment outcomes, including level of CU traits and affective empathy. Findings demonstrated a trend-level group difference ($p = .09$) in rate of change over time for CU traits, which although not significant, echoed the pattern seen for conduct problems in that children in the standard PCIT condition deteriorated during follow-up while the PCIT-CU condition showed sustained gains. In contrast, standard PCIT was unexpectedly superior to PCIT-CU in improving affective empathy during the active treatment period. While baseline group differences in CU traits and empathy make these statistical findings somewhat difficult to interpret, indicators of clinical significance demonstrated that only a small proportion of children were rated as showing reliable improvement in these outcomes from baseline to follow-up, irrespective of group. This is consistent with the only other study investigating improvement in CU traits following an adapted PMT program in a sample of CP + CU children, which found that mean levels of CU traits remained in the clinical range following the adapted treatment (Dadds et al., 2019).

Considered together, these findings highlight three important considerations for future research. First, they suggest that despite indications of statistically significant improvement, clinically significant CU traits may be less amenable to change via standard and adapted PMT than is currently supposed (e.g., Hawes et al., 2014). Second, they highlight the need for reliable, valid, and normed measures of affective outcomes to assess clinical normalization of CU traits and empathy following intervention. Finally, they demonstrate the importance of considering both between-group and within-participant differences when evaluating treatment efficacy, which existing studies of standard (e.g., Kimonis et al., 2014) and adapted

(e.g., Dadds et al., 2019) treatments for CP + CU populations have generally failed to do.

Finally, indicators of engagement demonstrated excellent acceptability of both standard PCIT and PCIT-CU, with no significant group differences in treatment attrition or satisfaction. We also demonstrated for the first time that treatment adherence and homework completion did not differ for standard compared to adapted treatment. These findings are somewhat inconsistent with prior PCIT research showing that CU traits predicted dropout (Kimonis et al., 2014) and low treatment satisfaction (Fleming et al., 2020). Although the dropout rate of 23% in this randomized controlled trial is at the lower end of the range reported in the literature for standard PCIT (i.e., 25–69%; Webb et al., 2017), addressing attrition remains a critical goal. Indeed, PCIT dropouts have significantly worse outcomes compared to treatment completers when assessed 1–3 years following termination (Boggs et al., 2005). In the broader PCIT literature, remotely delivered PCIT has been offered as a solution to the problem of attrition due to inadequate treatment accessibility (Comer et al., 2017). While evidence suggests that the detrimental effect of CU traits is particularly pronounced for Internet-delivered PCIT (Fleming et al., 2020), the results of an initial case study of online PCIT-CU were promising (Fleming et al., 2017).

Current findings must be considered within the context of several study limitations. First, this trial was conducted in a university research clinic within a relatively affluent region of Sydney, as reflected by the sociodemographic characteristics of the sample. As such, the performance of PCIT-CU was evaluated under ideal and highly controlled circumstances, likely overestimating the intervention's effect when implemented under conditions more closely resembling real-world clinical practice or within more sociodemographically diverse populations. On the other hand, we delivered PCIT-CU and standard PCIT using a fixed-length format instead of PCIT's typical approach of terminating treatment only once parents have achieved a prespecified level of skill proficiency and report that child behavior has normalized. Thus, our methodological approach may have limited parents' capacity to master skills and thus engender child behavioral change, possibly resulting in an underestimation of the effects of both standard PCIT and PCIT-CU. Related, families in standard PCIT received 7 weeks of consultation calls, which also represents a deviation from the standard PCIT protocol. While this may have enhanced the overall impact of standard PCIT rel-

ative to other studies, and relative to PCIT-CU participants since parents received continuous support in implementing PCIT strategies while CARES participants worked on emotional skills, the calls were not a true equivalent to the hour-long, in-person CARES sessions completed by PCIT-CU participants. Accordingly, we cannot rule out the possibility that treatment dosage or length of follow-up may explain the association between PCIT-CU and sustained treatment gains. Taken together, alongside further support for its efficacy, an effectiveness trial is needed to examine the performance of PCIT-CU compared to standard PCIT under such real-world conditions.

Second, group differences in observed and teacher-reported changes in behavioral and affective outcomes were not examined, limiting the generalizability of findings across settings and informants. Third, families were only followed up to 3 months posttreatment, limiting generalizability over time. Given that group differences were apparent only during the follow-up period, a longer follow-up is required to substantiate whether PCIT-CU confers sustained and clinically meaningful benefit over and above standard PCIT in the long term. Finally, the pilot nature of this proof-of-concept trial necessitated a small sample size that likely limited power to detect significant Group \times Time interaction effects, although inclusion of indicators of clinical significance goes some way in ameliorating this limitation. Nonetheless, several of the effects reported are only trending toward statistical significance, so circumspection is recommended when interpreting the findings. A trial with a considerably larger sample size is required to investigate whether the effects reported here are replicable and, crucially, to investigate mechanisms of change.

Limitations notwithstanding, this study has several key strengths. First, this is one of only two randomized trials to exclusively recruit a CP + CU sample of children in investigating outcomes associated with parent training intervention. Second, the targeted treatment was specifically designed to intervene during a critical period for conscience development, when CU traits can first be reliably identified. Finally, the targeted intervention was nested within an established treatment model. Because PCIT is a widely known and accepted treatment, PCIT-CU can be readily disseminated and implemented by the global network of PCIT practitioners. This rigorous study represents a timely and persuasive challenge to a long history of pessimism and stigma regarding the treatability of CU traits by showing that, at least early in development, these affective and interper-

sonal characteristics are amenable to treatment. It continues the important process of designing, testing, implementing, and disseminating comprehensive and individualized treatments for children with conduct problems and CU traits to enhance their response to such interventions.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.beth.2022.07.001>.

References

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA preschool/school-age forms & profiles*. University of Vermont Research Center for Children, Youth, & Families.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author. <http://doi.org/10.1176/appi.books.9780890425596>
- Barker, E. D., Oliver, B. R., Viding, E., Salekin, R. T., & Maughan, B. (2011). The impact of prenatal maternal risk, fearless temperament and early parenting on adolescent callous-unemotional traits: A 14-year longitudinal investigation. *Journal of Child Psychology and Psychiatry*, 52(8), 878–888. <https://doi.org/10.1111/j.1469-7610.2011.02397.x>.
- Bjørnebekk, G., & Kjøbli, J. (2017). Observed callousness as a predictor of treatment outcomes in parent management training. *Clinical Child Psychology and Psychiatry*, 22(1), 59–73. <https://doi.org/10.1177/1359104515621961>.
- Blizzard, A. M., Barroso, N. E., Ramos, F. G., Graziano, P. A., & Bagner, D. M. (2018). Behavioral parent training in infancy: What about the parent–infant relationship? *Journal of Clinical Child and Adolescent Psychology*, 47 (Suppl. 1), S341–S353. <https://doi.org/10.1080/15374416.2017.1310045>.
- Boggs, S. R., Eyberg, S. M., Edwards, D. L., Rayfield, A., Jacobs, J., Bagner, D., & Hood, K. K. (2005). Outcomes of parent–child interaction therapy: A comparison of treatment completers and study dropouts one to three years later. *Child and Family Behavior Therapy*, 26(4), 1–22. https://doi.org/10.1300/J019v26n04_01.
- Brestan, E. V., Jacobs, J. R., Rayfield, A. D., & Eyberg, S. M. (1999). A consumer satisfaction measure for parent–child treatments and its relation to measures of child behavior change. *Behavior Therapy*, 30, 17–30. [https://doi.org/10.1016/S0005-7894\(99\)80043-4](https://doi.org/10.1016/S0005-7894(99)80043-4).
- Byrd, A. L., Loeber, R., & Pardini, D. A. (2014). Antisocial behavior, psychopathic features and abnormalities in reward and punishment processing in youth. *Clinical Child and Family Psychology Review*, 17(2), 125–156. <https://doi.org/10.1007/s10567-013-0159-6>.
- Carter Leno, V., Charman, T., Pickles, A., Jones, C. R. G., Baird, G., Happe, F., & Simonoff, E. (2015). Callous-unemotional traits in adolescents with autism spectrum disorder. *British Journal of Psychiatry*, 207, 392–399. <https://doi.org/10.1192/bjp.bp.114.159863>.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>.
- Comer, J. S., Furr, J. M., Miguel, E. M., Cooper-Vince, C. E., Carpenter, A. L., Elkins, R. M., Kerns, C. E., Cornacchio, D., Chou, T., Cox, S., DeSerisy, M., Sanchez, A. L., Golik,

- A., Martin, J., Myers, K. M., & Chase, R. (2017). Remotely delivering real-time parent training to the home: An initial randomized trial of Internet-delivered parent-child interaction therapy (I-PCIT). *Journal of Consulting and Clinical Psychology, 85*(9), 909–917. <https://doi.org/10.1037/ccp0000230>.
- Dadds, M. R., Cauchi, A. J., Wimalaweera, S., Hawes, D. J., & Brennan, J. (2012). Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Research, 199*(3), 201–207. <https://doi.org/10.1016/j.psychres.2012.04.033>.
- Dadds, M. R., El Masry, Y., Wimalaweera, S., & Guastella, A. J. (2008). Reduced eye gaze explains “fear blindness” in childhood psychopathic traits. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(4), 455–463. <https://doi.org/10.1097/CHI.0b013e31816407f1>.
- Dadds, M. R., English, T., Wimalaweera, S., Schollar-Root, O., & Hawes, D. J. (2019). Can reciprocated parent-child eye gaze and emotional engagement enhance treatment for children with conduct problems and callous-unemotional traits: A proof-of-concept trial. *Journal of Child Psychology and Psychiatry, 60*(6), 676–685. <https://doi.org/10.1111/jcpp.13023>.
- Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D., Vassallo, S., Bunn, P., Merz, S., & El Masry, Y. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development, 39*, 111–122. <https://doi.org/10.1007/s10578-007-0075-4>.
- Datyner, A., Kimonis, E. R., Hunt, E., & Armstrong, K. (2016). Using a novel emotional skills module to enhance empathic responding for a child with conduct disorder with limited prosocial emotions. *Clinical Case Studies, 15*(1), 35–52. <https://doi.org/10.1177/1534650115588978>.
- Dawel, A., O’Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews, 36*, 2288–2304. <https://doi.org/10.1016/j.neubiorev.2012.08.006>.
- Doig, G. S., & Simpson, F. (2005). Randomization and allocation concealment: A practical guide for researchers. *Journal of Critical Care, 20*(2), 187–191. <https://doi.org/10.1016/j.jcrc.2005.04.005>.
- Dunn, J., Brown, J. R., & Maguire, M. (1995). The development of children’s moral sensibility: Individual differences and emotion understanding. *Developmental Psychology, 31*(4), 649–659. <https://doi.org/10.1037/0012-1649.31.4.649>.
- Ellis, H. D. (1990). Development trends in face recognition. *The Psychologist, 3*, 114–119.
- Eyberg, S. M., & Funderburk, B. (2011). *Parent-child interaction therapy protocol*. PCIT International.
- Eyberg, S. M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory—Revised: Professional manual*. PAR.
- Ezpeleta, L., Osa, N. D. L., Granero, R., Penelo, E., & Domènech, J. M. (2013). Inventory of Callous-Unemotional Traits in a community sample of preschoolers. *Journal of Clinical Child and Adolescent Psychology, 42*(1), 91–105. <https://doi.org/10.1080/15374416.2012.734221>.
- Feingold, A. (2019). Time-varying effect sizes for quadratic growth models in multilevel and latent growth modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(3), 418–429. <https://doi.org/10.1080/10705511.2018.1547110>.
- Fleming, G. E., & Kimonis, E. R. (2018). PCIT for children with callous-unemotional traits. In L. N. Niec (Ed.), *Handbook of parent-child interaction therapy* (pp. 19–34). Springer. https://doi.org/10.1007/978-3-319-97698-3_2.
- Fleming, G. E., Kimonis, E. R., Datyner, A., & Comer, J. S. (2017). Adapting Internet-delivered parent-child interaction therapy to treat co-occurring disruptive behavior and callous-unemotional traits: A case study. *Clinical Case Studies, 16*(5), 370–387. <https://doi.org/10.1177/1534650117699471>.
- Fleming, G. E., Kimonis, E. R., Furr, J. M., & Comer, J. S. (2020). Internet-delivered parent training for preschoolers with conduct problems: Do callous-unemotional traits moderate efficacy and engagement? *Journal of Abnormal Child Psychology, 48*(9), 1169–1182. <https://doi.org/10.1007/s10802-020-00660-5>.
- Frick, P. J., & Kemp, E. C. (2021). Conduct disorders and empathy development. *Annual Review of Clinical Psychology, 17*, 391–416. <https://doi.org/10.1146/annurev-clinpsy-081219-105809>.
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014a). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry, 55*(6), 532–548. <https://doi.org/10.1111/jcpp.12152>.
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014b). 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>.
- Garcia, A. M., Graziano, P. A., & Hart, K. C. (2018). Response to time-out among preschoolers with externalizing behavior problems: The role of callous-unemotional traits. *Child Psychiatry and Human Development, 49*(5), 699–708. <https://doi.org/10.1007/s10578-018-0788-6>.
- 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*(4), 737–741. <https://doi.org/10.1037/0022-006X.73.4.737>.
- 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*(3), 248–267. <https://doi.org/10.1007/s10567-014-0167-1>.
- Holland, D., & Dadds, M. R. (1997). *The Diagnostic Interview Schedule for Children, Adolescents, and Parents*. Griffith University.
- Hyde, L. W., Shaw, D. S., Gardner, F., Cheong, J., Dishion, T. J., & Wilson, M. (2013). Dimensions of callousness in early childhood: Links to problem behavior and family intervention effectiveness. *Development and Psychopathology, 25*(2), 347–363. <https://doi.org/10.1017/S0954579412001101>.
- Hyde, L. W., Waller, R., Trentacosta, C. J., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Reiss, D., & Leve, L. D. (2016). Heritable and nonheritable pathways to early callous-unemotional behaviors. *American Journal of Psychiatry, 173*(9), 903–910. <https://doi.org/10.1176/appi.ajp.2016.15111381>.

- 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*(1), 12–19. <https://doi.org/10.1037/10109-042>.
- Kimonis, E. R., Bagner, D. M., Linares, D., Blake, C. A., & Rodriguez, G. (2014). Parent-training outcomes among young children with callous-unemotional conduct problems with or at-risk for developmental delay. *Journal of Child and Family Studies, 23*, 437–448. <https://doi.org/10.1007/s10826-013-9756-8>.
- Kimonis, E. R., Fanti, K. A., Anastassiou-Hadjicharalambous, X., Mertan, B., Goulter, N., & Katsimicha, E. (2016). Can callous-unemotional traits be reliably measured in preschoolers? *Journal of Abnormal Child Psychology, 44*(4), 625–638. <https://doi.org/10.1007/s10802-015-0075-y>.
- Kimonis, E. R., Fanti, K. A., Frick, P. J., Moffitt, T. E., Essau, C., Bijttebier, P., & Marsee, M. A. (2015). Using self-reported callous-unemotional traits to cross-nationally assess the DSM-5 “with limited prosocial emotions” specifier. *Journal of Child Psychology and Psychiatry, 56*(11), 1249–1261. <https://doi.org/10.1111/jcpp.12357>.
- Kimonis, E. R., Fleming, G. E., Briggs, N., Brouwer-French, L., Frick, P. J., Hawes, D. J., Bagner, D. M., Thomas, R., & Dadds, M. R. (2019). Parent-child interaction therapy adapted for preschoolers with callous-unemotional traits: An open trial pilot study. *Journal of Clinical Child and Adolescent Psychology, 48*(Suppl. 1), S347–S361. <https://doi.org/10.1080/15374416.2018.1479966>.
- Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology, 33*(2), 228–240.
- 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*(4), 713–725. <https://doi.org/10.1037/a0020910>.
- Moore, A. A., Blair, R. J., Hettema, J. M., & Roberson-Nay, R. (2019). The genetic underpinnings of callous-unemotional traits: A systematic research review. *Neuroscience and Biobehavioral Reviews, 100*, 85–97. <https://doi.org/10.1016/j.neubiorev.2019.02.018>.
- Morrissey-Kane, E., & Prinz, R. J. (1999). Engagement in child and adolescent treatment: The role of parental cognitions and attributions. *Clinical Child and Family Psychology Review, 2*(3), 183–198. <https://doi.org/10.1023/A:1021807106455>.
- Moul, C., Hawes, D. J., & Dadds, M. R. (2018). Mapping the developmental pathways of child conduct problems through the neurobiology of empathy. *Neuroscience and Biobehavioral Reviews, 91*, 34–50. <https://doi.org/10.1016/j.neubiorev.2017.03.016>.
- Moul, C., Killcross, S., & Dadds, M. R. (2012). A model of differential amygdala activation in psychopathy. *Psychological Review, 119*, 789–806. <https://doi.org/10.1037/a0029342>.
- Palm, S. M., Sawrikar, V., Schollar-Root, O., Moss, A., Hawes, D. J., & Dadds, M. R. (2019). Parents’ spontaneous attributions about their problem child: Associations with parental mental health and child conduct problems. *Journal of Abnormal Child Psychology, 47*(9), 1455–1466. <https://doi.org/10.1007/s10802-019-00536-3>.
- Pardini, D. A., Lochman, J. E., & Powell, N. (2007). The development of callous-unemotional traits and antisocial behavior in children: Are there shared and/or unique predictors? *Journal of Clinical Child and Adolescent Psychology, 36*(3), 319–333. <https://doi.org/10.1080/15374410701444215>.
- Pasalich, D. S., Dadds, M. R., Hawes, D. J., & Brennan, J. (2011). Do callous-unemotional traits moderate the relative importance of parental coercion versus warmth in child conduct problems? An observational study. *Journal of Child Psychology and Psychiatry, 52*(12), 1308–1315. <https://doi.org/10.1111/j.1469-7610.2011.02435.x>.
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry, 56*, 345–365. <https://doi.org/10.1111/jcpp.12381>.
- Sawrikar, V., & Dadds, M. (2018). What role for parental attributions in parenting interventions for child conduct problems? Advances from research into practice. *Clinical Child and Family Psychology Review, 21*(1), 41–56. <https://doi.org/10.1007/s10567-017-0243-4>.
- Thomas, R., & Zimmer-Gembeck, M. J. (2012). Parent-child interaction therapy: An evidence-based treatment for child maltreatment. *Child Maltreatment, 17*(3), 253–266. <https://doi.org/10.1177/1077559512459555>.
- 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*(4), 593–608. <https://doi.org/10.1016/j.cpr.2013.03.001>.
- Waller, R., Gardner, F., Shaw, D. S., Dishion, T. J., Wilson, M. N., & Hyde, L. W. (2015). Callous-unemotional behavior and early-childhood onset of behavior problems: The role of parental harshness and warmth. *Journal of Clinical Child and Adolescent Psychology, 44*(4), 655–667. <https://doi.org/10.1080/15374416.2014.886252>.
- Waller, R., Hyde, L. W., Grabbell, A. S., Alves, M. L., & Olson, S. L. (2015). Differential associations of early callous-unemotional, oppositional, and ADHD behaviors: Multiple domains within early-starting conduct problems? *Journal of Child Psychology and Psychiatry, 56*(6), 657–666. <https://doi.org/10.1111/jcpp.12326>.
- Webb, H. J., Thomas, R., McGregor, L., Avdagic, E., & Zimmer-Gembeck, M. J. (2017). An evaluation of parent-child interaction therapy with and without motivational enhancement to reduce attrition. *Journal of Clinical Child and Adolescent Psychology, 46*(4), 537–550. <https://doi.org/10.1080/15374416.2016.1247357>.
- World Health Organization (2019). *International statistical classification of diseases and related health problems* (11th ed.). Author. Retrieved from <https://icd.who.int>.
- Wright, N., Pickles, A., Sharp, H., & Hill, J. (2021). A psychometric and validity study of callous-unemotional traits in 2.5 year old children. *Scientific Reports, 11*(8065), 1–10. <https://doi.org/10.1038/s41598-021-87416-3>.

RECEIVED: December 6, 2021

ACCEPTED: July 7, 2022

AVAILABLE ONLINE: XXXX