CAMBRIDGE
UNIVERSITY PRESS

Regular Article

Associations of observed callous–unemotional behaviors in early childhood with conduct problems and substance use over 14 years

Meghan Rose Donohue [iD], Caroline P. Hoyniak, Rebecca Tillman, Deanna M. Barch and Joan Luby

Department of Psychiatry, Washington University School of Medicine, Saint Louis, USA

Abstract

Callous–unemotional (CU) behaviors (i.e., low concern and active disregard for others) uniquely predict severe conduct problems and substance use when present by late childhood. Less is known about the predictive utility of CU behaviors displayed in early childhood, when morality is developing and interventions may be more effective. Children aged 4–7 years ($N = 246$; 47.6% girls) completed an observational task wherein they were encouraged to tear an experimenter's valued photograph, and blind raters coded children's displayed CU behaviors. During the next 14 years, children's conduct problems (i.e., oppositional defiant and conduct symptoms) and age of onset of substance use were assessed. Compared to children displaying fewer CU behaviors, children displaying greater CU behaviors were 7.61 times more likely to meet criteria for a conduct disorder ($n = 52$) into early adulthood (95% CI, 2.96–19.59; $p = <.0001$), and their conduct problems were significantly more severe. Greater CU behaviors were associated with earlier onset of substance use ($B = −.69$, $SE = .32$, $t = −2.14$, $p = .036$). An ecologically valid observed indicator of early CU behavior was associated with substantially heightened risk for conduct problems and earlier onset substance use into adulthood. Early CU behavior is a powerful risk marker identifiable using a simple behavioral task which could be used to target children for early intervention.

Keywords: callous–unemotional, conduct problems, early childhood, observational coding, substance use

(Received 4 December 2020; revised 14 June 2021; accepted 18 June 2021; First published online 9 August 2021)

Severe conduct problems (i.e., oppositional defiant and conduct disorder symptoms) are associated with increased risk for criminality and substance use and often result in substantial harm to victims, perpetrators, and society (Bongers, Koot, Van Der Ende, & Verhulst, 2004; Foster & Jones, 2005; Odgers et al., 2008b). As such, identifying at-risk children is of critical public health importance. To date, studies have identified "early-starting" conduct problems (i.e., those that onset by age 5 years) as one of the most robust predictors of later severe conduct problems (Loeber & Dishion, 1983; Shaw & Gross, 2008). Yet, as a substantial proportion of early-starting children do not continue to display conduct problems into adolescence or adulthood (Maughan & Rutter, 1998; Odgers et al., 2008b), research identifying factors that improve upon the longitudinal prediction of severe conduct problems and their associated poor outcomes is greatly needed.

One potentially robust predictor of later severe conduct problems and substance use is the presence of callous–unemotional (CU) behaviors. CU behavior is a multidimensional construct encompassing low concern (i.e., low empathy, guilt, and prosocial behavior) and active disregard for others (i.e., taking pleasure in others' distress; Rhee et al., 2013; Waller et al., 2020) and thought to be a developmental precursor to psychopathy (Frick, 1995). Longitudinal studies have demonstrated that CU behaviors displayed in late childhood and adolescence add to the prediction of subsequent conduct problems and substance use beyond the most well-established predictors of these outcomes, including children's baseline conduct problem severity and whether their problems were early starting (Frick, Ray, Thornton, & Kahn, 2014). For example, 7- to 11-year-olds who displayed greater CU behaviors also exhibited more severe conduct problems 4 years later, controlling for their baseline conduct problem severity (Pardini, Byrd, Hawes, & Docherty, 2018).

Further, greater CU behaviors in adolescence have been found to predict earlier and more frequent substance use (Anderson, Zheng, & McMahon, 2018; Baskin-Sommers, Waller, Fish, & Hyde, 2015; Ray, Thornton, Frick, Steinberg, & Cauffman, 2016; Thornton et al., 2019; Waller & Hicks, 2019). The fact that adolescent CU behaviors predicted earlier substance use is particularly notable given that early-onset use of alcohol and drugs is one of the strongest predictors of later substance use disorders and is associated with outcomes such as early pregnancy, school dropout, and criminality (DeWit, Adlaf, Offord, & Ogborne, 2000; Odgers et al., 2008a). Compared to individuals with conduct problems without CU behaviors, there is a unique, etiological pathway that is thought to contribute to risk for early-onset substance use disorders in individuals with CU behaviors (Brennan, Hyde, & Baskin-Sommers, 2017). In this pathway,

CrossMark

novelty/thrill-seeking behaviors, heightened sensitivity to the rewards associated with substance use, and decreased sensitivity to – and thereby decreased capacity to learn from – the negative consequences of substance use, put individuals with CU behaviors at heightened risk for substance use (Brennan et al., 2017). This theoretical model, along with growing evidence of the association between adolescent CU behaviors and later substance use, highlights the importance of exploring early CU behaviors as predictors of later substance use.

Despite evidence of associations between CU behaviors displayed in late childhood and adolescence and later conduct problems and substance use, less is known about the predictive value of CU behaviors displayed during early childhood – the developmental period when emerging capacities for empathy, guilt, and prosocial behavior rapidly develop and may be most modifiable (Kochanska, Gross, Lin, & Nichols, 2002; Roth-Hanania, Davidov, & Zahn-Waxler, 2011; Waller & Hyde, 2018). Research on the typical development of the moral emotions and behaviors that develop atypically in children with greater CU behaviors has consistently demonstrated that children's empathic and prosocial tendencies emerge during infancy and develop rapidly across the first years of life (Davidov et al., 2021; Hay, Nash, & Pedersen, 1981; Roth-Hanania et al., 2011; Schuhmacher, Collard, & Kärtner, 2017; Sommerville, Schmidt, Yun, & Burns, 2013). Existing examinations of early childhood CU behaviors have either measured outcomes at only one time point (Dadds, Fraser, Frost, & Hawes, 2005; Ezpeleta, Osa, Granero, Penelo, & Domènech, 2013, 2015; Jezior, McKenzie, & Lee, 2016; López-Romero, Romero, & Luengo, 2012; Moran et al., 2009; Rowe, Costello, Angold, Copeland, & Maughan, 2010; Waller et al., 2014, 2015, 2016) and/or examined narrow follow-up periods (e.g., 1–3 years; Pardini, Obradovic, & Loeber, 2006; Pardini & Fite, 2010), potentially missing symptoms occurring between time points or during later developmental periods. Although there is evidence that early childhood CU behaviors predict aggression and rule breaking in middle childhood (Waller et al., 2016), whether early CU behaviors predict more distal outcomes remains unknown. Further, no study has examined whether CU behaviors displayed during early childhood are predictive of later problematic substance use. Establishing early childhood CU behaviors as a unique long-term predictor of conduct problems and problematic substance use could inform the identification of at-risk children and development of targeted preventative intervention. These efforts are of utmost importance, as treatments may be most efficacious in early childhood, a period of potentially heightened neuroplasticity (Johnson, 2005).

As the study of early childhood CU behaviors is in its infancy, developmentally appropriate measurement methods have been limited. To date, most studies have used parent reports of both early CU behaviors and subsequent outcomes (Longman, Hawes, & Kohlhoff, 2016), which has been found to inflate resulting associations (Waller et al., 2016). To surmount this limitation, the present study used an observed indicator of CU behaviors measured during an *in vivo* task to predict parent reports and self-reports of later conduct and substance use problems. Although observational measures have not been widely used to measure CU behaviors, such tasks are a well-validated means through which developmentalists have studied typical moral development for decades (Kochanska et al., 2002; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Such observational measures could be an important source of information about a child's moral emotions and behaviors in addition to parent and/or teacher report and may be a particularly important method through which to assess CU behaviors, as moral emotions can be highly internal.

One of the few studies to observationally measure components of the CU behaviors construct across four time points at 14, 20, 24, and 36 months of age found that observed active disregard but not low empathy and prosocial behavior predicted children's greater subsequent conduct problems in adolescence across parent, teacher, and child informants (Rhee et al., 2013). Moreover, a follow-up study of this sample found that observed active disregard but not low empathy and prosocial behavior across toddlerhood predicted greater interpersonal and affective symptoms of psychopathy in adulthood (Rhee et al., 2020). This body of work supports the construct validity and substantial potential predictive utility of early observed measures of aberrant morality, and suggests that active disregard may be particularly predictive of subsequent conduct problems.

Thus, although there is promising evidence that CU behaviors in childhood may predict later severe conduct problems and problematic substance use, it is unclear whether CU behaviors displayed in early childhood add prognostic value beyond other established predictors. Identifying early childhood CU behavior as a unique predictor of conduct and substance use outcomes into early adulthood would provide compelling evidence of the clinical utility of burgeoning efforts to measure CU behaviors this early in development. The present prospective longitudinal study examined observations of CU behaviors in early childhood as a potential predictor of conduct problems (measured prospectively from childhood through early adulthood) and age of substance use onset. We hypothesized that greater CU behaviors in early childhood would uniquely predict subsequent conduct problem diagnoses and severity, and earlier onset of substance use.

Methods

Participants

Participants included 246 children from the preschool depression study conducted at Washington Universitiy School of Medicine. Informed consent and assent were obtained from participants and all procedures were approved by the institutional review board. At the baseline assessment (Time 1 or T1), 3- to 6-year-old children with elevated symptoms of depression and externalizing disorders (attention-deficit/hyperactivity disorder [ADHD] and/or oppositional defiant disorder [ODD]) and children without symptoms of psychiatric disorders (i.e., healthy children) were recruited. A score of 3 or higher on the preschool feelings checklist (Luby, Heffelfinger, Koenig-McNaught, Brown, & Spitznagel, 2004) was used to identify children with elevated symptoms of depression and externalizing disorders, whose parents then completed diagnostic interviews at the baseline assessment. This study examines the 246 participants who completed the picture tearing task at T3. Of these 246 participants, $n = 54$ met criteria for depression, $n = 70$ met criteria for another psychiatric disorder, and $n = 119$ were healthy at the baseline assessment. During the next 17 years, children completed nine follow-up assessments. The current investigation examines data from all 10 study time points, through T10, which occurred when children were 15 to 21 years old (see Figure 1). Note that diagnostic status tended to be fluid over the course of the 17 years in this longitudinal observational cohort study; for example, of the $n = 119$ children who were healthy at the study baseline, $n = 84$ went on to meet criteria for a psychiatric disorder at a subsequent study time point. Nonetheless, children completed a similar number of
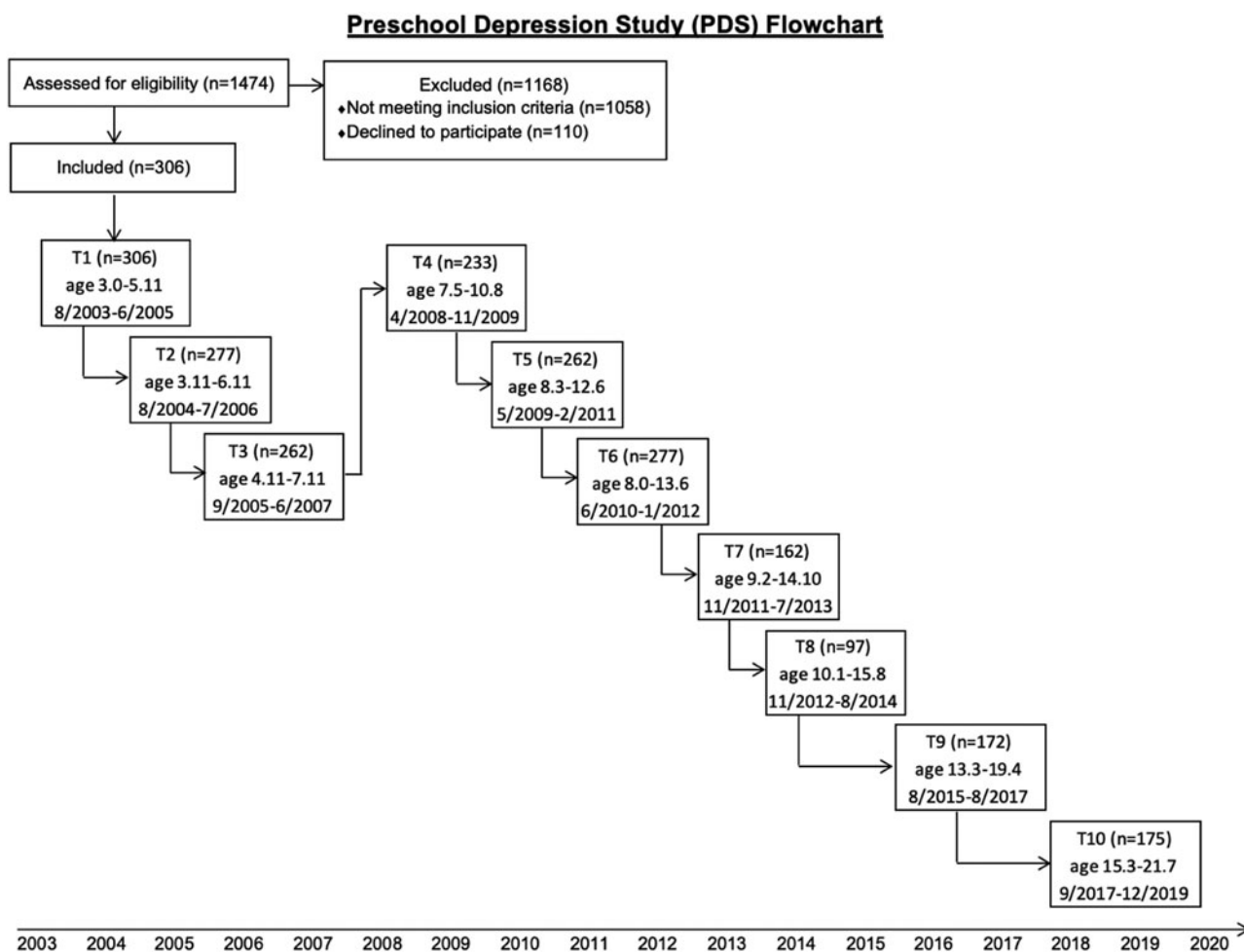
Preschool Depression Study (PDS) Flowchart

Figure 1. Flowchart depicting time points of the preschool depression study (PDS).

total assessment waves regardless of their diagnostic group membership at baseline (e.g., on average, between 7.10 and 7.48 total waves).

Measures

Observed CU behaviors

Observations of CU behaviors, this study's primary predictor variable, were videotaped and assessed at T3, when children were 4 to 7 years old, via the picture tearing task, part of the laboratory temperament assessment battery (Lab-TAB; Goldsmith, Reilly, Lemery, Longley, & Prescott, 2001). Although the task was originally designed to measure children's compliance with a directive, it also elicits emotions and behaviors following a transgression and is thus ideal for observational measurement of CU behaviors. This task has been used successfully to measure children's shame and inappropriate/excessive guilt; in this study, fathers' depression and permissive parenting and marital dissatisfaction when children were age 3 years predicted greater expressions of guilt and shame at age 6 years (Parisette-Sparks, Bufferd, & Klein, 2017).

In the task, an experimenter (the "victim") first shows the child a photograph of her grandparents, and explains to the child it is her "favorite" photograph because they live far away and it helps her remember them. The victim leaves the room and another experimenter ("E") enters and, using a neutral tone, directs the child to tear the photograph, saying, "Now (child's name), I want you to

take (victim name's) favorite picture out of the album and tear it up." If the child does not tear the photograph right away, E waits up to 2 min (m) for a response, prompting at specific intervals: specifically, at the 60 seconds (s) and 90 s mark, E restates the initial prompt, saying, "Go ahead, tear up (victim name's) favorite picture." If the child asks questions about why E wants them to tear the photograph, E responds, "I just want you to." If the child becomes too distressed at any time, or if after 2 m the child still has not torn the photograph, the paradigm is stopped, and E says, "That wasn't such a good idea anyway. Let's do something else." If the child tears the photo, the victim returns, asks the child, "What happened?" and sits with the child while displaying mild distress for 1 min; no further prompts are used. This period gives the child the opportunity to spontaneously display moral emotions and behaviors toward the victim after transgressing. Finally, the child is debriefed: E pretends to leave the room to see if there is a copy of the photograph, re-enters, shows the child an intact copy and says, "It probably wasn't such a good idea when I told you to tear up (victim name's) favorite picture. I'm glad she had a copy of her picture. Let's play a game." In sum, as an unfamiliar adult asks the child to commit a transgression that would be hurtful to another adult, the task presents a unique challenge, setting up a conflict between the need to comply with a directive and any internal resistance to engaging in a hurtful act against another.

Research assistants (RAs) blind to hypotheses and children's clinical characteristics rated displays of CU behaviors during the task. Each of the four components of the CU behaviors construct was coded on a scale ranging from 0 = *none* to 3 = *strong*. Note that children's displays of moral emotions and behaviors were entirely unprompted. For example, children were not prompted to use prosocial behavior (e.g., to fix the photo or generate ideas about how to fix the photo). Coders rated children's displays of the following two variables using the coding scheme from the published Lab-TAB manual.

Empathy
Coders rated children's peak intensity of empathy/concern verbalization *prior* to tearing the photo (or refusing to tear the photo). This coding most closely reflected children's expressions of the cognitive component empathy, which involves attempts to apprehend another's emotional state, specifically information seeking (that is, attempts to understand another's mental state and viewpoint; (Eisenberg & Fabes, 1998; Shamay-Tsoory, 2009). Empathy was coded using the following scale: 0 = *no empathic or concerned verbalizations*; 1 = *child repeats the directions back for E to clarify or directly asks for clarification of the directions* (e.g., to make sure the child tester really wants him/her to violate the norm); 2 = *child asks questions about why E wants him/her to break the norm, or asks if the owner of the picture will be mad or sad if the picture is torn up*; 3 = *child shows serious empathic concern about violating the norm*. Child directly states (not in the form of a question) something about the owner of the picture and the awareness that s/he may be unhappy if the child follows through with the task.

Active disregard
Coders rated children's behavioral and affective enjoyment in tearing up the photo on the scale ranging from 0 = *none* to 3 = *strong* as a measure of active disregard for the victim. Behavioral expressions of enjoyment included tearing the photo excessively (e.g., tearing the photo into tiny pieces, continuing to tear the photo long after given the initial instruction, continuing to tear the photo in the presence of the victim). Affective expressions of enjoyment while tearing the photo included smiling, laughing, and other expressions of positive affect.

The final two variables were rated on a scale also ranging from 0 = *none* to 3 = *strong* using a coding scheme adapted from that previously published by one of the authors (Donohue, Williamson, & Tully, 2020). Raters considered the frequency, duration, and intensity/sophistication of affect/behaviors displayed after tearing the photo when rating the following:

Guilt affect
Coders rated evidence of children's guilt affect after tearing the photo following the definition of guilt, which is an emotional and cognitive experience triggered by a transgression that involves two components: experiencing affective empathy – negative emotion in response to another's emotional distress – and an awareness of personal responsibility for causing the transgression (Tilghman-Osborne, Cole, & Felton, 2010). Guilt also involves focus on the wrongdoing itself (I did that bad *thing*; Tracy & Robins, 2004). Thus, coders rated children's vocal, facial, and postural displays of other-oriented empathy toward the victim after tearing the photo, indications that children understood they were personally responsible for the transgression (e.g., statements of regret, such as "I wish I didn't do it"), and cognitive and affective focus on the mishap.

Prosocial behavior
Coders rated children's displays of prosocial behaviors – actions that are helpful, supportive, or intended to benefit another person – after tearing the photo. Common displays of prosocial behaviors during this task included attempts to repair the photo (e.g., piecing together the torn pieces), statements about repairing the photo (e.g., making suggestions about how to repair the photo), and comforting statements (e.g., "You're okay;" "I'll help you"). Other examples of prosocial behaviors coded included apologies ("I'm sorry"), confessions (e.g., "I ripped your picture") and information seeking questions (e.g., "Are you sad?").

Ratings of empathy, guilt, and prosocial behavior were each reverse scored to yield measures of *low empathy*, *low guilt*, and *low prosociality*. An observed CU behaviors composite score was created by averaging the three reverse-scored variables and the *active disregard* variable. A principal components analysis demonstrated that the composite scale represented a one-factor solution: there was only one eigenvalue greater than 1 (1.68) with all item loadings above .30 ($M = .63$). In the current sample, children with greater observed CU behaviors also demonstrated lower levels of parent-reported empathy, prosocial behavior, and guilt across T1–T3, demonstrating construct validity for this observed measure (see Supplementary Material 1, Table S1, Appendix). Inter-rater reliability was assessed for a random 25% of videos with intraclass correlation coefficients; agreement was in the excellent range ($M_{ICC} = .83$).

Psychiatric diagnoses and severity
Diagnostic interviews were conducted at each time point that assess the presence and severity of psychiatric disorders using the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) criteria. The interview utilized depended on children's age. The preschool age psychiatric assessment (Egger, Angold, Small, &Copeland, 1999) was administered to parents when participants were younger than 8 years old, the child and adolescent psychiatric assessment (CAPA; Angold & Costello, 2000) to parents and children when participants were 8–17 years old, and the kiddie-schedule for affective disorders and schizophrenia (K-SADS; Kaufman et al., 1997) to parents and adolescents when participants were 18 years or older.

This study primarily examined the presence and severity of conduct disorders, including ODD and conduct disorder (CD). Severity was assessed by summing the total number of core symptoms endorsed for a given disorder. Other disorders were examined as covariates or in specificity analyses. Children were considered to meet criteria for a symptom if endorsed by either parent or child. The presence and severity of nonconduct psychiatric disorders including ADHD, anxiety disorders (i.e., generalized anxiety disorder, separation anxiety disorder, and social phobia), and major depressive disorder (MDD) were examined either as covariates or in specificity analyses. Whereas diagnostic information was available through T10, severity scores were available through T8 only due to differences in the CAPA and K-SADS. Satisfactory inter-rater reliabilities for these diagnostic interviews in this sample have been previously reported (Whalen, Belden, Tillman, Barch, & Luby, 2016).

Substance use
Substance use was assessed at T10 with the composite international diagnostic interview for the fourth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) Version 2.0 (Robins et al., 1988), a standardized instrument for assessing

substance use disorders. The interview assesses use of the following substances: alcohol, tobacco, marijuana, sedatives, stimulants, painkillers, cocaine, club drugs, hallucinogens, opioids, inhalants, and other drugs. In the current study, problematic substance use was indexed by age of onset of substance use, measured by averaging the self-reported age of first use of each of the above substances.

Data analytic plan

Analyses examined whether children's observed CU behaviors predicted subsequent conduct problem diagnosis and severity and earlier onset substance use over and above well-established predictors of these outcomes: early-starting conduct problems and baseline symptom severity. All hypotheses tested were made a priori based on prior literature. First, logistic regression was used to determine whether CU behaviors at T3 (age 4–7 years) predicted odds of meeting criteria for conduct disorders at one or more time points from T4 to T10 (age 7.5–21.5 years). Each model controlled for child sex, child age at the T3 assessment, income-to-needs ratio at T3, mean severity across T1–T3 of ADHD and anxiety, whether the child had preschool-onset MDD (PO-MDD; defined as MDD that emerged before age 6 years), and whether the child had early-starting conduct problems (i.e., met criteria for ODD or CD by T3). Second, dimensional analyses used multilevel modeling to determine whether CU behaviors at T3 predicted conduct problem severity across T4–T9. The multilevel models included a random intercept with all other variables fixed effects; time was coded as age at assessment wave. Finally, multiple regression was used to examine whether CU behaviors at T3 predicted children's age of onset of substance use. Multilevel and regression models controlled for child sex, income-to-needs, PO-MDD diagnosis, and mean severity across T1–T3 of ADHD, anxiety, and conduct problems. The regression model additionally controlled for children's age at T3. False discovery rate (FDR) correction for multiple comparisons was used.

Results

Given the research questions, children were excluded from analysis of study hypotheses if they did not tear the photo (i.e., "nontearers;" $n = 57$). Children were also excluded if they did not complete at least one assessment after T3 ($n = 18$), yielding a final sample of $n = 171$. Tearers had significantly lower income-to-needs ratios than nontearers (Table 1). A significant difference in the proportion of tearers by race was no longer significant when controlling for income-to-needs ($p = .64$). There was some evidence that tearers exhibited more concurrent and subsequent conduct problems and less observed and parent-reported empathy than nontearers, although some of these associations were only marginally significant. Nontearers are likely to be a small but heterogeneous group composed of some children that did not feel comfortable committing a transgression as well as other children who may have refused to comply with the directive to tear the photo out of oppositionality/defiance rather than moral objection. Children who tore the photo exhibited significantly greater conduct disorder severity across T1–T3 compared to children who did not tear the photo. Similarly, tearers were more likely to develop later CD/ODD than nontearers, but this association was only marginally significant (Supplementary Table S2). Tearers ($M = 1.28$, $SD = .79$) displayed significantly less observed (coded) empathy prior to tearing the photo than nontearers ($M = 2.14$, $SD = .93$), t (244) = 6.88,

$p = <.001$. Tearers ($M = 4.82$, $SD = .55$) also displayed less parent-reported empathy, guilt, and prosocial behavior than nontearers ($M = 4.96$, $SD = .49$), but this difference was only marginally significant, t (243) = 1.77, $p = .08$. Analyses in-text report on the subsample of tearers only ($n = 171$), unless otherwise indicated.

There was a high degree of psychiatric comorbidity in the sample. Across T1–T3, 24.6% met criteria for both a conduct problem diagnosis (i.e., ODD or CD) and an internalizing disorder diagnosis (i.e., MDD or an anxiety disorder), 24.6% met criteria for an internalizing disorder only, and 12.3% met criteria for a conduct disorder only; 38.6% of the sample did not meet criteria for any internalizing or conduct disorder.

Regarding conduct problem diagnoses specifically, $n = 63$ (36.8%) of children met criteria for ODD or CD by T3, with an average age of onset of 4.66 years (SD = .85), indicating a high proportion of children with early-starting conduct problems in the sample. Of this group of early-starting children, 41.27% remitted in that they never again met criteria for ODD/CD after T3. Across T4–T10, $n = 52$ (30.4%) of children in the full sample of tearers met criteria for ODD/CD at one or more time points.

Correlations among variables are presented in Table 2. Older children and males displayed significantly greater CU behaviors. Observed CU behaviors were not significantly associated with concurrent severity of depression, externalizing problems, or anxiety. Observed CU behaviors were associated with higher odds of concurrent ODD/CD diagnoses, but this association was only marginally significant (Table S3 in the Supplementary Material).

Observed CU behaviors predicting subsequent conduct problems and substance use onset

Analyses first examined whether children's observed CU behaviors predicted subsequent conduct problem diagnoses, beyond whether children had early-starting conduct problems. Children's greater observed CU behaviors in early childhood were uniquely associated with 7.61 times higher odds of subsequently meeting criteria for ODD or CD into early adulthood when controlling for children's concurrent and prior ODD and CD diagnoses (Table 3; Figure 2). Breaking this down into risk for each disorder, children's greater CU behaviors were associated with 9.42 times higher odds of subsequently meeting criteria for ODD and 5.75 times higher odds of subsequently meeting criteria for CD. Within the subgroup of children with early-starting conduct problems, greater CU behaviors significantly predicted conduct problem persistence (i.e., continuing to meet criteria for ODD/CD at any time after T3; Supplementary Table S4). Using a sum score of CU behaviors (i.e., summing scores across the four subcomponents), a supplemental ROC analysis revealed that a cut-off score of 7 yielded the best trade-off in sensitivity and specificity in predicting which children were most likely to subsequently meet criteria for ODD or CD (Supplementary Figure S2).

Dimensionally, children's greater observed CU behaviors in early childhood uniquely predicted significantly greater conduct disorder severity from middle childhood through early adulthood, beyond concurrent and prior conduct disorder severity and when examining ODD and CD both combined and separately (Table 4). Observed CU behaviors showed specificity in predicting conduct problem outcomes, as they were not significantly associated with odds of subsequently meeting criteria for ADHD, an anxiety disorder, or MDD, or the subsequent severity of these disorders (see Supplementary Material, Tables S5 and S6). Furthermore, greater

Table 1. Descriptive statistics

Variable	Observed range	Mean (*SD*) or frequency		χ^2 or *t* value[b]	*p* value
		Tearers[a] (*n*=189)	Nontearers (*n*=57)		
1. T3 age (years)	4.99–7.99	6.49 (.81)	6.42 (.71)	−.60	.55
2. Sex (% female)		47.6%	52.6%	.44	.51
3. Race (%)				11.37	.003
White		50.8%	68.4%		
Black		39.7%	15.8%		
Bi/multiracial		9.5%	15.8%		
4. Income-to-needs[c]	0–4.38	1.90 (1.08)	2.42 (.93)	3.29	.001
T3					
1. Depression severity	0–8	2.03 (1.76)	2.23 (1.81)	.75	.45
2. Externalizing severity	0–27	5.63 (6.09)	5.82 (5.22)	.21	.83
3. ODD severity	0–7	1.43 (1.70)	1.58 (1.53)	.59	.55
4. CD severity	0–7	.92 (1.18)	.71 (.99)	−1.21	.23
5. Internalizing severity	0–14	1.65 (2.48)	1.93 (2.91)	.71	.48
Mean across T1–T3					
1. ODD severity	0–7.3	1.77 (1.61)	1.67 (1.27)	−.50	.62
2. CD severity	**0–5.5**	**1.02 (1.14)**	**.66 (.72)**	**−2.85**	**.005**
T4–T10	Observed range	Tearers[a] (*n* = 171)	Nontearers (*n*=52)	χ^2 or *t* value[b]	*p* value
1. ODD diagnosis (% present)		26.3%	17.3%	1.76	.18
2. CD diagnosis (% present)		17.5%	7.7%	2.99	.08
3. ODD severity	0–7	1.45 (1.50)	1.21 (1.51)	−.99	.32
4. CD severity	0–6	.68 (.93)	.39 (.59)	−2.61	.01
5. Age of substance use onset	11.7–19.0	16.44 (1.36)	16.78 (.99)	1.06	.29

Bold is to indicate that the analysis was significant (*p* < .05).
[a]Tearers = children who tore the photo; nontearers = children who did not tear the photo
[b]Independent samples *t* tests were used to compare continuous variables across groups, χ^2 tests to compare categorical/binary variables across groups, and Fisher's exact tests (F.E.) when there were small expected cell counts.
[c]Families' income-to-needs ratio, or the family's income divided by the federal poverty level based on family size, was included as a covariate (McLoyd, 1998); a ratio of 1 signifies that the child's family was living at the poverty threshold.

Table 2. Correlations among variables in subsample of children who tore photo (*n* = 189)

Variable	1	2	3	4	5	6	7
1. T3 age							
2. Female sex	−.13						
3. T3 income to needs	.17*	−.01					
4. **T3 observed CU behaviors**	.20**	−.17*	−.02				
5. T3 depression severity	.13	−.06	−.12	.03			
6. T3 externalizing severity	.02	−.13	−.19*	.02	.57**		
7. T3 anxiety severity	.09	.00	.07	−.00	.48**	.35**	

*=*p* < .05; **=*p* < .01

observed CU behaviors significantly predicted earlier substance use onset (Table 5) when controlling for a range of psychiatric comorbidities. Sensitivity analyses were conducted that included both tearers and nontearers to examine whether effects of CU behaviors predicting subsequent conduct disorder diagnoses and severity and substance use held; nontearers were assigned the highest score for guilt and prosociality and the lowest score for active disregard to be as conservative as possible. All effects

Table 3. Logistic regression models of subsequent oppositional defiant disorder/conduct disorder (ODD/CD) diagnoses by observed callous–unemotional (CU) behaviors

	B	SE	Wald	p	FDR p	Exp. (B)	95% CI for exp (B)
DV: CD/ODD (T4–T10)							
Intercept	−5.54	1.97	7.88	.005			
Female sex	−.03	.22	.01	.91		.95	.40–2.28
T3 age	.13	.28	.22	.64		1.14	.66–2.00
T3 income-to-needs	−.03	.20	.02	.89		.97	.65–1.45
T1–T3 ADHD severity	.07	.07	1.28	.26		1.08	.95–1.23
T1–T3 anxiety severity	.08	.13	.41	.52		1.09	.84–1.41
PO-MDD diagnosis	.17	.25	.46	.50		1.41	.52–3.81
T1–T3 ODD/CD diagnosis	1.13	.26	18.96	<.0001		9.67	3.48–26.85
T3 observed CU behaviors	**2.03**	**.48**	**17.70**	**<.0001**	**.0002**	**7.61**	**2.96–19.59**
DV: ODD T4-10							
Intercept	−4.81	2.01	5.69	.02			
Female sex	−.19	.24	.64	.42		.69	.27–1.73
T3 age	−.10	.30	.13	.72		.90	.50–1.61
T3 income-to-needs	.16	.22	.58	.45		1.18	.77–1.80
T1–T3 ADHD severity	.11	.07	2.37	.12		1.11	.97–1.27
T1–T3 anxiety severity	.01	.14	.01	.92		1.01	.78–1.32
PO-MDD diagnosis	.50	.26	3.65	.06		2.72	.97–7.62
T1–T3 ODD diagnosis	1.06	.27	15.15	<.0001		8.40	2.88–24.54
T3 observed CU behaviors	**2.24**	**.51**	**19.11**	**<.0001**	**.0002**	**9.42**	**3.45–25.76**
DV: CD T4-T10							
Intercept	−6.66	2.59	6.62	.01			
Female sex	.16	.31	.27	.60		1.38	.42–4.56
T3 age	.32	.36	.78	.38		1.38	.68–2.81
T3 income-to-needs	−.59	.28	4.38	.04		.55	.32–.96
T1–T3 ADHD severity	.11	.07	2.17	.14		1.12	.96–1.29
T1–T3 anxiety severity	.24	.15	2.43	.12		1.27	.94–1.72
PO-MDD diagnosis	.05	.31	.03	.87		1.11	.32–3.80
T1–T3 CD diagnosis	1.14	.28	16.08	<.0001		9.71	3.20–29.49
T3 observed CU behaviors	**1.75**	**.62**	**7.91**	**.005**	**.005**	**5.75**	**1.70–19.43**

Note: ADHD, attention-deficit/hyperactivity disorder; FDR, false discovery rate; ODD/CD, oppositional defiant disorder/conduct disorder; PO-MDD, preschool onset major depressive disorder.

continued to be significant in these analyses (see Supplementary Material, Tables S7 and S8). As older children and males displayed greater CU behaviors, age and gender were tested as moderators of the significant associations between CU behaviors and conduct problems and substance use; none of these models was significant (see Supplementary Material, Tables S9–S14).

Examining the four individual components of the CU behaviors composite score in the same model demonstrated that low empathy, low prosociality, and active disregard each uniquely predicted subsequent severity of ODD and CD symptoms, with low prosociality being the most predictive component (i.e., largest effect size; see Supplementary Material, Table S15). Quite similarly, subsequent severity of ODD alone was also significantly predicted by low empathy, low prosociality, and active disregard, with low prosociality the most predictive (see Supplementary Material, Table S16). In contrast, subsequent CD severity and age of onset of substance use were not significantly predicted by any individual component (see Supplementary Material, Tables S17 and S18).

Discussion

These longitudinal findings are the first to demonstrate that observed CU behaviors in early childhood robustly predict subsequent conduct and substance use problems into early adulthood. Findings could inform identification of children at high risk for conduct and substance use problems and underscore that elevated CU behaviors may be a prime intervention target in early childhood, while morality is developing.
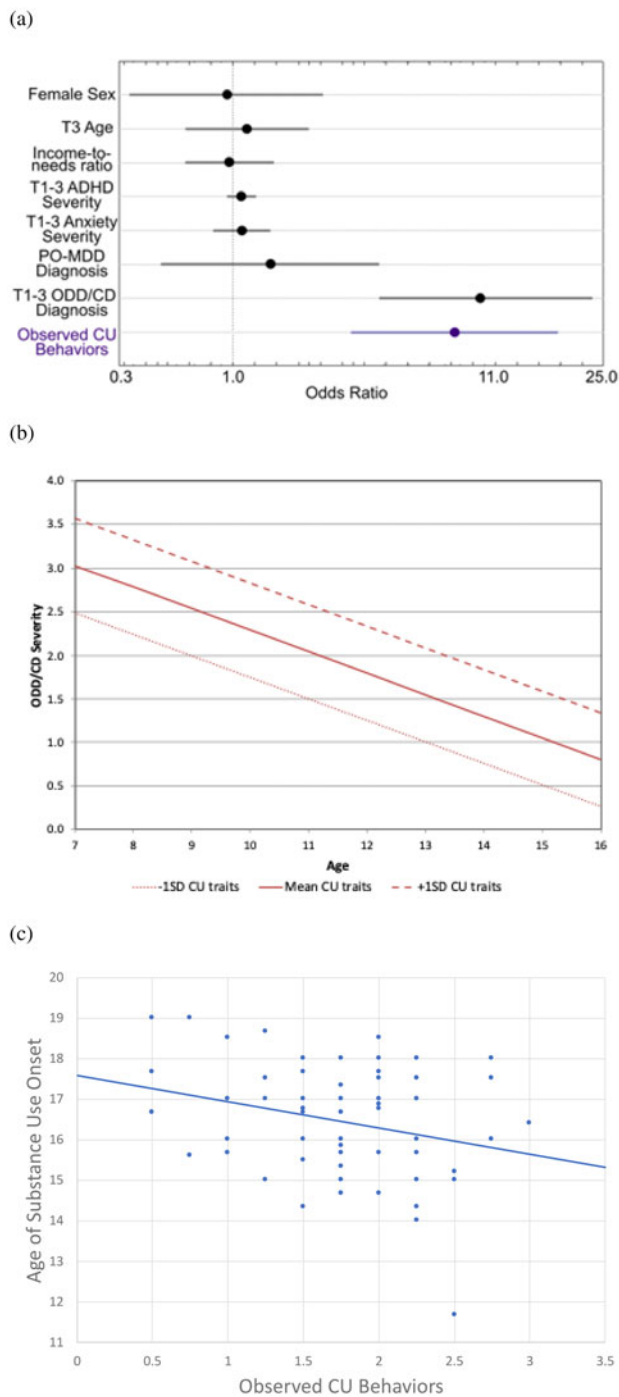
(a)

(b)

(c)

Figure 2. (a) Forest plot depicting observed callous–unemotional (CU) behaviors predicting subsequent odds of meeting criteria for oppositional defiant disorder/conduct disorder (ODD/CD) and line graphs depicting observed CU behaviors as a predictor of subsequent (b) severity of ODD/CD and (c) age of substance use onset.

Children who displayed greater CU behaviors during the observational task during early childhood were more than seven times more likely subsequently to develop a conduct disorder into early adulthood than children who displayed fewer CU behaviors, and their conduct problems were more severe. Moreover, observed CU behaviors were unique predictors of conduct disorder outcomes beyond two of the most well-established risk factors – the presence of early-starting conduct problems

and baseline conduct problem severity. Greater observed CU behaviors specifically predicted conduct problems rather than other psychiatric disorders. Greater observed CU behaviors also predicted whether children's early-starting conduct problems persisted. Given that the conduct problems of nearly half of early-starting children in this sample remitted, observations of CU behaviors may meaningfully inform assessments of risk for enduring conduct problems. Thus, observations of young children's CU behaviors demonstrated added value to the prediction of very distal conduct problem outcomes. Findings extend previous research indicating that early childhood CU behaviors predict conduct problems into middle childhood (Waller et al., 2016) by demonstrating that this prediction extends into early adulthood. This evidence represents some of the strongest support to date for the predictive utility of observationally measuring CU behaviors this early in development.

Our findings support prior work, which found that observed measures of aberrant morality in a community sample of toddlers predicted greater conduct problems in adolescence (Rhee et al., 2013, 2020), and extends these findings in a high-risk, early childhood sample. Interestingly, whereas in these prior studies, observed active disregard but not low empathy or prosociality predicted subsequent conduct problems, in the current study, all three of these components predicted subsequent conduct problems, with low prosociality being the most predictive individual component. Thus, our findings are somewhat in contrast to prior work that suggested active disregard is more predictive of later antisocial behavior than more "passive" empathy deficits such as low empathy and prosociality (Rhee et al., 2013, 2020), a difference that is possibly due to distinctions in the observational tasks used between the studies or the fact that the current study examines a clinical sample whereas the studies by Rhee et al. were conducted in community samples. Yet our finding that low prosociality was the most predictive individual component of the CU behaviors composite is in line with recent meta-analytic work demonstrating that, of low empathy, prosociality, and guilt, low prosociality shows the most robust association with CU traits (Waller et al., 2020). Our findings contribute to converging evidence of the importance of low prosociality in understanding CU behaviors and outcomes.

Greater CU behaviors observed in early childhood also predicted an earlier onset of alcohol and drug use. Several correlates of CU behaviors are thought to put individuals at risk for problematic substance use including low sensitivity to punishment and empathic deficits. For example, individuals with low empathy may be less likely to care about how their substance use affects their family and peers, promoting substance use onset and maintenance (Massey, Newmark, & Wakschlag, 2018). Prior studies have demonstrated that CU behaviors displayed in late childhood and adolescence were associated with more problematic adolescent substance use (Anderson et al., 2018; Muratori et al., 2016; Ray et al., 2016). Our study is the first to our knowledge to demonstrate that as early as age four, children displaying high levels of CU behaviors appear to be at risk for initiating substance use at younger ages.

In this study, observations of CU behaviors were behaviorally coded from an ecologically valid task in which children are directed by an adult to hurt another person and are subsequently faced with the victim they have harmed. In contrast to previous studies, our coding scheme comprehensively measured all four components of the CU behavior construct (Waller et al., 2020), including low levels of empathy, guilt, and prosociality, and active disregard. Importantly, the task demonstrated initial long-term

Table 4. Multilevel models of observed callous–unemotional (CU) behaviors predicting the severity of conduct problems

Outcome	B	SE	t	p	FDR p
ODD and CD severity					
Intercept	−0.805	0.915	−0.88	0.3802	
Assessment wave	−0.321	0.055	−5.88	<0.0001	
Female sex	0.240	0.215	1.12	0.2649	
T3 age	−0.046	0.135	−0.35	0.7305	
T3 income-to-needs	0.015	0.099	0.1	0.8788	
T1–T3 ODD/CD severity	0.449	0.060	7.52	<0.0001	
T1–T3 ADHD severity	0.120	0.039	3.10	0.0023	
T1–T3 anxiety severity	0.040	0.069	0.57	0.5680	
PO-MDD diagnosis	0266	0.257	1.04	0.3010	
T3 observed CU behaviors	**0.985**	**0.200**	**4.93**	**<0.0001**	**<0.0001**
ODD severity					
Intercept	0099	0.679	0.15	0.8840	
Assessment wave	−0.230	0.040	−5.79	<0.0001	
Female sex	0.062	0.160	0.39	0.6994	
T3 age	−0.157	0.100	−1.57	0.1194	
T3 income-to-needs	0.126	0.074	1.71	0.0886	
T1–T3 ODD/CD severity	0.266	0.044	6.01	<0.0001	
T1–T3 ADHD severity	0.080	0.029	2.79	0.0060	
T1–T3 anxiety severity	−0.014	0.051	−0.28	0.7790	
PO-MDD diagnosis	0.385	0.191	2.02	0.0452	
T3 observed CU behaviors	**0.744**	**0.149**	**5.00**	**<0.0001**	**<0.0001**
CD severity					
Intercept	−1.007	0.407	−2.48	0.0143	
Assessment wave	−0.087	0.025	−3.51	0.0006	
Female sex	0.188	0.095	1.97	0.0511	
T3 age	0.122	0.060	2.04	0.0427	
T3 income-to-needs	−0.107	0.044	−2.44	0.0159	
T1–T3 ODD/CD severity	0.189	0.027	7.11	<0.0001	
T1–T3 ADHD severity	0.037	0.017	2.16	0.0324	
T1–T3 anxiety severity	0.052	0.031	1.68	0.0941	
PO-MDD diagnosis	−0.116	0.114	−1.02	0.3095	
T3 observed CU behaviors	**0.243**	**0.089**	**2.73**	**0.0071**	**0.0071**

Note: ADHD, attention-deficit/hyperactivity disorder; FDR, false discovery rate; ODD/CD, oppositional defiant disorder/conduct disorder; PO-MDD, preschool onset major depressive disorder.

validity in predicting later conduct problems and substance use. This brief, cost-efficient task and straightforward coding scheme could easily be implemented in clinical or school settings for targeted children to provide another avenue to assess CU behaviors beyond parent or teacher reports. Moreover, the use of an observed measure circumvents the potential biases of relying only on parent reports of CU behaviors. Of note, whereas greater CU behaviors were highly predictive of *subsequent* conduct problems, they were not significantly associated with *current* problems. This suggests that the task could identify high-risk children that would otherwise be missed through assessments that rely on diagnostic information alone. Preliminary analyses suggest that a sum score of 7 or higher was useful in identifying children most likely to develop later conduct problems in our sample; further replication in community populations is essential, however, before this score would be ready for use in clinical settings – identifying such a cut-off score could eventually greatly aid clinicians in identifying the children at greatest risk for later conduct problems to refer for early intervention. Moreover, our finding that low prosociality was the most predictive individual component of CU behavior could similarly assist teachers, caregivers, and clinicians in pinpointing which children are most at risk for later problems, and in directly and proactively assessing these behaviors in the classroom or community.

Table 5. Multiple regression model of observed callous–unemotional (CU) behaviors predicting the age of onset of substance use

	B	SE	t	p	FDR p
Substance use onset					
Intercept	15.28	1.59	9.60	<.0001	
Female sex	−.33	.35	−.92	.36	
T3 age	.43	.20	2.15	.04	
T3 income-to-needs	.06	.15	.38	.71	
T1–T3 ODD/CD severity	−.13	.09	−1.49	.14	
T1–T3 ADHD	−.02	.05	−.30	.76	
T1–T3 anxiety severity	.01	.10	.08	.94	
PO-MDD diagnosis	−.06	.37	−.14	.89	
Observed CU behaviors	**−.69**	**.32**	**−2.14**	**.036**	**.036**

Note: ADHD, attention-deficit/hyperactivity disorder; FDR, false discovery rate; ODD/CD, oppositional defiant disorder/conduct disorder; PO-MDD, preschool onset major depressive disorder.

Of note, this study's sample was enriched for preschool depression, although only a subset of participants met criteria for preschool depression and the sample contained a comparable proportion of children with externalizing disorders. Whereas a growing literature supports the existence of two CU subtypes – primary and secondary CU, marked by low and high levels of anxiety, respectively (Fanti, Demetriou, & Kimonis, 2013; Kimonis, Frick, Cauffman, Goldweber, & Skeem, 2012) – much less research has focused on CU behaviors in the context of depression. To date, studies examining associations between CU behaviors and depression are mixed, with some evidence suggesting that CU behaviors may be related to less depression (Pardini & Fite, 2010), and other evidence of greater depressive symptoms in children with elevated CU traits (Craig & Moretti, 2019; Salekin, Leistico, Neumann, DiCicco, & Duros, 2004). Future studies should examine whether the current findings generalize to young children displaying higher levels of CU behaviors without comorbid depressive symptoms. On the one hand, it is unclear whether findings would generalize to children with clinical levels of conduct problems and very elevated levels of CU traits. On the other hand, findings may not be unique to children with depression given the high level of comorbidity between internalizing and externalizing disorders, in our sample and in the literature generally.

This study had a number of other strengths, including its prospective design and mixed-method and multiple informant measurement approach, which improves upon prior research that has largely relied on parent report alone (Waller et al., 2016). Moreover, our observed measure comprehensively assessed all components of the CU behavior construct and was significantly associated with parent report of morality. However, it would have been ideal to include a parent-report measure of CU behaviors; specifically, future studies should examine both parent-reported and observed CU behaviors to compare the predictive utility of both measurement approaches. Our study also has other limitations, including our measure of age of onset of substance use, which was based on self-report at one time point, and thus included retrospective report. Future studies should assess substance use at multiple time points.

Observed CU behaviors in early childhood uniquely and powerfully predicted conduct problem diagnoses and severity into early adulthood and earlier substance use onset. Findings demonstrate

that assessing early CU behaviors using highly feasible observational tasks may markedly improve upon the identification of children at high risk for conduct and substance use problems. Further, findings underscore the importance of treating CU behaviors in early childhood. Recent research has identified treatment approaches, including family-focused interventions and modifications of traditional Parent × Child interaction therapy, that target core features of CU behaviors such as emotion understanding deficits and are effective at reducing CU behaviors in young children (Donohue, Hoyniak, TIllman, Barch, & Luby, in press; Hyde et al., 2013; Kimonis et al., 2019). Our findings underscore the importance of implementing such interventions in early childhood, when CU behaviors are more malleable, in order to alter CU behavior trajectories and prevent antisocial behavior into adulthood.

References

Anderson, S. L., Zheng, Y., & McMahon, R. J. (2018). Do callous–unemotional traits and conduct disorder symptoms predict the onset and development of adolescent substance Use? *Child Psychiatry & Human Development*, 49, 688–698. doi:10.1007/s10578-018-0789-5

Angold, A., & Costello, E. J. (2000). The child and adolescent psychiatric assessment (CAPA). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 39–48. doi:10.1097/00004583-200001000-00015

Baskin-Sommers, A. R., Waller, R., Fish, A. M., & Hyde, L. W. (2015). Callous-unemotional traits trajectories interact with earlier conduct problems and executive control to predict violence and substance use among high risk male adolescents. *Journal of Abnormal Child Psychology*, 43, 1529–1541. doi:10.1007/s10802-015-0041-8

Bongers, I. L., Koot, H. M., Van Der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, 75, 1523–1537.

Brennan, G. M., Hyde, L. W., & Baskin-Sommers, A. R. (2017). Antisocial pathways associated with substance use disorders: Characterizing etiological

underpinnings and implications for treatment. *Current Opinion in Behavioral Sciences*, 13, 124–129. doi:10.1016/j.cobeha.2016.11.014

Craig, S. G., & Moretti, M. M. (2019). Profiles of primary and secondary callous-unemotional features in youth: The role of emotion regulation. *Development and Psychopathology*, 31, 1489–1500. doi:10.1017/S0954579418001062

Dadds, M. R., Fraser, J., Frost, A., & Hawes, D. J. (2005). Disentangling the underlying dimensions of psychopathy and conduct problems in childhood: A community study. *Journal of Consulting and Clinical Psychology*, 73, 400–410. doi:10.1037/0022-006X.73.3.400

Davidov, M., Paz, Y., Roth-Hanania, R., Uzefovsky, F., Orlitsky, T., Mankuta, D., & Zahn-Waxler, C. (2021). Caring babies: Concern for others in distress during infancy. *Developmental Science*, 24, e13016.

DeWit, D. J., Adlaf, E. M., Offord, D. R., & Ogborne, A. C. (2000). Age at first alcohol use: A risk factor for the development of alcohol disorders. *American Journal of Psychiatry*, 157, 745–750.

Donohue, M. R., Hoyniak, C. P., TIllman, R., Barch, D. M., & Luby, J. (in press). Callous-unemotional traits as an intervention target and moderator of PCIT-ED treatment for preschool depression and conduct problems. *Journal of the American Academy of Child & Adolescent Psychiatry*.

Donohue, M. R., Williamson, R. A., & Tully, E. C. (2020). Toddlers imitate prosocial demonstrations in bystander but not transgressor contexts. *Journal of Experimental Child Psychology*, 192, 104776.

Egger, H. L., Angold, A., Small, B., & Copeland, W. (1999). The preschool age psychiatric assessment. In *The Oxford Handbook of Infant, Toddler, and Preschool Mental Health Assessment*.

Eisenberg, N., & Fabes, R. A. (1998). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development*, Vol. 3, 5th ed. (pp. 701–778). Hoboken, NJ: John Wiley & Sons Inc.

Ezpeleta, L., Granero, R., de la Osa, N., & Domènech, J. M. (2015). Clinical characteristics of preschool children with oppositional defiant disorder and callous-unemotional traits. *PLoS One*, 10, e0139346. doi:10.1371/journal.pone.0139346

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 & Adolescent Psychology*, 42, 91–105. doi:10.1080/15374416.2012.734221

Fanti, K. A., Demetriou, C. A., & Kimonis, E. R. (2013). Variants of callous-unemotional conduct problems in a community sample of adolescents. *Journal of Youth and Adolescence*, 42, 964–979.

Foster, E. M., Jones, D. E., & Conduct Problems Prevention Research Group. (2005). The high costs of aggression: Public expenditures resulting from conduct disorder. *American Journal of Public Health*, 95, 1767–1772.

Frick, P. J. (1995). Callous-unemotional traits and conduct problems: A two-factor model of psychopathy in children. *Issues in Criminological & Legal Psychology*, 24, 47–51.

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

Goldsmith, H., Reilly, J., Lemery, K., Longley, S., & Prescott, A. (2001). The laboratory temperament assessment battery: Middle childhood version. Unpublished manuscript, University of Wisconsin-Madison.

Hay, D. F., Nash, A., & Pedersen, J. (1981). Responses of six-month-olds to the distress of their peers. *Child Development*, 52, 1071–1075. doi:10.2307/1129114

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, 347–363. doi:10.1017/S0954579412001101

Jezior, K. L., McKenzie, M. E., & Lee, S. S. (2016). Narcissism and callous-unemotional traits prospectively predict child conduct problems. *Journal of Clinical Child & Adolescent Psychology*, 45, 579–590. doi:10.1080/15374416.2014.982280

Johnson, M. H. (2005). Sensitive periods in functional brain development: Problems and prospects. *Developmental Psychobiology*, 46, 287–292. doi:10.1002/dev.20057

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children – present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36, 980–988. doi:10.1097/00004583-199707000-00021

Kimonis, E. R., Fleming, G., Briggs, N., Brouwer-French, L., Frick, P. J., Hawes, D. J., … Dadds, M. (2019). Parent-child interaction therapy adapted for preschoolers with callous-unemotional traits: An open trial pilot study. *Journal of Clinical Child & Adolescent Psychology*, 48, S347–S361. doi:10.1080/15374416.2018.1479966

Kimonis, E. R., Frick, P. J., Cauffman, E., Goldweber, A., & Skeem, J. (2012). Primary and secondary variants of juvenile psychopathy differ in emotional processing. *Development and Psychopathology*, 24, 1091–1103. doi:10.1017/S0954579412000557

Kochanska, G., Gross, J. N., Lin, M.-H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, 73, 461–482. doi:10.1111/1467-8624.00418

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin*, 94, 68.

Longman, T., Hawes, D. J., & Kohlhoff, J. (2016). Callous–unemotional traits as markers for conduct problem severity in early childhood: A meta-analysis. *Child Psychiatry & Human Development*, 47, 326–334. doi:10.1007/s10578-015-0564-9

López-Romero, L., Romero, E., & Luengo, MÁ. (2012). Disentangling the role of psychopathic traits and externalizing behaviour in predicting conduct problems from childhood to adolescence. *Journal of Youth and Adolescence*, 41, 1397–1408. doi:10.1007/s10964-012-9800-9

Luby, J. L., Heffelfinger, A., Koenig-McNaught, A. L., Brown, K., & Spitznagel, E. (2004). The preschool feelings checklist: A brief and sensitive screening measure for depression in young children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43, 708–717. doi:10.1097/01.chi.0000121066.29744.08

Massey, S. H., Newmark, R. L., & Wakschlag, L. S. (2018). Explicating the role of empathic processes in substance use disorders: A conceptual framework and research agenda. *Drug and Alcohol Review*, 37, 316–332.

Maughan, B., & Rutter, M. (1998). Continuities and discontinuities in antisocial behavior from childhood to adult life. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology*. Advances in Clinical Child Psychology, vol 20. Boston, MA: Springer. https://doi-org.libproxy.wustl.edu/10.1007/978-1-4757-9038-2_1

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist*, 53(2), 185–204.

Moran, P., Rowe, R., Flach, C., Briskman, J., Ford, T., Maughan, B., … Goodman, R. (2009). Predictive value of callous-unemotional traits in a large community sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48, 1079–1084. doi:10.1097/CHI.0b013e3181b766ab

Muratori, P., Lochman, J. E., Manfredi, A., Milone, A., Nocentini, A., Pisano, S., & Masi, G. (2016). Callous unemotional traits in children with disruptive behavior disorder: Predictors of developmental trajectories and adolescent outcomes. *Psychiatry Research*, 236, 35–41. doi:10.1016/j.psychres.2016.01.003

Odgers, C. L., Caspi, A., Nagin, D. S., Piquero, A. R., Slutske, W. S., Milne, B. J., … Moffitt, T. E. (2008a). Is it important to prevent early exposure to drugs and alcohol among adolescents? *Psychological Science*, 19, 1037–1044.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2008b). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, 20, 673–716. doi:10.1017/S0954579408000333

Pardini, D. A., Byrd, A. L., Hawes, S. W., & Docherty, M. (2018). Unique dispositional precursors to early-onset conduct problems and criminal offending in adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57, 583–592.e3. doi:10.1016/j.jaac.2018.04.013

Pardini, D. A., & Fite, P. J. (2010). Symptoms of conduct disorder, oppositional defiant disorder, attention-deficit/hyperactivity disorder, and callous-unemotional traits as unique predictors of psychosocial maladjustment in boys: Advancing an evidence base for DSM-V. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49, 1134–1144. doi:10.1016/j.jaac.2010.07.010

Pardini, D., Obradovic, J., & Loeber, R. (2006). Interpersonal callousness, hyperactivity/impulsivity, inattention, and conduct problems as precursors to delinquency persistence in boys: A comparison of three grade-based cohorts. *Journal of Clinical Child & Adolescent Psychology*, 35, 46–59. doi:10.1207/s15374424jccp3501_5

Parisette-Sparks, A., Bufferd, S. J., & Klein, D. N. (2017). Parental predictors of children's shame and guilt at age 6 in a multimethod, longitudinal study. *Journal of Clinical Child & Adolescent Psychology*, 46, 721–731. doi:10.1080/15374416.2015.1063430

Ray, J. V., Thornton, L. C., Frick, P. J., Steinberg, L., & Cauffman, E. (2016). Impulse control and callous-unemotional traits distinguish patterns of delinquency and substance use in justice involved adolescents: Examining the moderating role of neighborhood context. *Journal of Abnormal Child Psychology*, 44, 599–611. doi:10.1007/s10802-015-0057-0

Rhee, S. H., Friedman, N. P., Boeldt, D. L., Corley, R. P., Hewitt, J. K., Knafo, A., … Zahn-Waxler, C. (2013). Early concern and disregard for others as predictors of antisocial behavior: Early concern and disregard for others. *Journal of Child Psychology and Psychiatry*, 54, 157–166. doi:10.1111/j.1469-7610.2012.02574.x

Rhee, S. H., Woodward, K., Corley, R. P., du Pont, A., Friedman, N. P., Hewitt, J. K., … Zahn-Waxler, C. (2020). The association between toddlerhood empathy deficits and antisocial personality disorder symptoms and psychopathy in adulthood. *Development and Psychopathology*, 1–11. doi:10.1017/S0954579419001676

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., … Regier, D. A. (1988). The composite international diagnostic interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, 45, 1069–1077.

Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, 34, 447–458. doi:10.1016/j.infbeh.2011.04.007

Rowe, R., Costello, E. J., Angold, A., Copeland, W. E., & Maughan, B. (2010). Developmental pathways in oppositional defiant disorder and conduct disorder. *Journal of Abnormal Psychology*, 119, 726–738. doi:10.1037/a0020798

Salekin, R. T., Leistico, A.-M. R., Neumann, C. S., DiCicco, T. M., & Duros, R. L. (2004). Psychopathy and comorbidity in a young offender sample: Taking a closer look at psychopathy's potential importance over disruptive behavior disorders. *Journal of Abnormal Psychology*, 113, 416–427. doi:10.1037/0021-843X.113.3.416

Schuhmacher, N., Collard, J., & Kärtner, J. (2017). The differential role of parenting, peers, and temperament for explaining interindividual differences in 18-months-olds' comforting and helping. *Infant Behavior and Development*, 46, 124–134. doi:10.1016/j.infbeh.2017.01.002

Shamay-Tsoory, S. G. (2009). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety, & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 215–232). Cambridge, MA: MIT Press. doi:10.7551/mitpress/9780262012973.003.0017

Shaw, D. S., & Gross, H. E. (2008). What we have learned about early childhood and the development of delinquency. In A. M. Liberman (Ed.), *The long view of crime: A synthesis of longitudinal research* (pp. 79–127). New York, NY: Springer.

Sommerville, J. A., Schmidt, M. F., Yun, J., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy*, 18, 40–66.

Thornton, L. C., Frick, P. J., Ray, J. V., Wall Myers, T. D., Steinberg, L., & Cauffman, E. (2019). Risky sex, drugs, sensation seeking, and callous unemotional traits in justice-involved male adolescents. *Journal of Clinical Child & Adolescent Psychology*, 48, 68–79. doi:10.1080/15374416.2017.1399398

Tilghman-Osborne, C., Cole, D. A., & Felton, J. W. (2010). Definition and measurement of guilt: Implications for clinical research and practice. *Clinical Psychology Review*, 30, 536–546. doi:10.1016/j.cpr.2010.03.007

Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, 15, 103–125.

Waller, R., Dishion, T. J., Shaw, D. S., Gardner, F., Wilson, M. N., & Hyde, L. W. (2016). Does early childhood callous-unemotional behavior uniquely predict behavior problems or callous-unemotional behavior in late childhood? *Developmental Psychology*, 52, 1805–1819. doi:10.1037/dev0000165

Waller, R., Gardner, F., Viding, E., Shaw, D. S., Dishion, T. J., Wilson, M. N., & Hyde, L. W. (2014). Bidirectional associations between parental warmth, callous unemotional behavior, and behavior problems in high-risk preschoolers. *Journal of Abnormal Child Psychology*, 42, 1275–1285. doi:10.1007/s10802-014-9871-z

Waller, R., & Hicks, B. M. (2019). Trajectories of alcohol and marijuana use among primary versus secondary psychopathy variants within an adjudicated adolescent male sample. *Personality Disorders: Theory, Research, and Treatment*, 10, 87–96. doi:10.1037/per0000303

Waller, R., & Hyde, L. W. (2018). Callous-unemotional behaviors in early childhood: The development of empathy and prosociality gone awry. *Current Opinion in Psychology*, 20, 11–16. doi:10.1016/j.copsyc.2017.07.037

Waller, R., Hyde, L. W., Grabell, 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, 657–666. doi:10.1111/jcpp.12326

Waller, R., Wagner, N. J., Barstead, M. G., Subar, A., Petersen, J. L., Hyde, J. S., & Hyde, L. W. (2020). A meta-analysis of the associations between callous-unemotional traits and empathy, prosociality, and guilt. *Clinical Psychology Review*, 75, 101809. doi:10.1016/j.cpr.2019.101809

Whalen, D. J., Belden, A. C., Tillman, R., Barch, D. M., & Luby, J. L. (2016). Early adversity, psychopathology, and latent class profiles of global physical health from preschool through early adolescence. *Psychosomatic Medicine*, 78, 1008–1018. doi:10.1097/PSY.0000000000000398

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28, 126–136.