Psychological Medicine

cambridge.org/psm

Original Article

Cite this article: Hawes SW, Waller R, Thompson WK, Hyde LW, Byrd AL, Burt SA, Klump KL, Gonzalez R (2020). Assessing callous-unemotional traits: development of a brief, reliable measure in a large and diverse sample of preadolescent youth. *Psychological Medicine* **50**, 456–464. https://doi.org/10.1017/S0033291719000278

Received: 16 July 2018
Revised: 19 December 2018
Accepted: 29 January 2019
First published online: 8 March 2019

Key words:
Antisocial behaviors; callous-unemotional traits; conduct disorder

Author for correspondence:
Samuel W. Hawes, E-mail: shawes@fiu.edu

CAMBRIDGE UNIVERSITY PRESS

Assessing callous-unemotional traits: development of a brief, reliable measure in a large and diverse sample of preadolescent youth

Samuel W. Hawes[1,2], Rebecca Waller[3], Wesley K. Thompson[4], Luke W. Hyde[5,6,7], Amy L. Byrd[8], S. Alexandra Burt[9], Kelly L. Klump[9] and Raul Gonzalez[1,2]

[1]Center for Children and Families, Florida International University, Miami, FL, USA; [2]Department of Psychology, Florida International University, Miami, FL, USA; [3]Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA; [4]Division of Biostatistics, Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA; [5]Department of Psychology, University of Michigan, Ann Arbor, MI, USA; [6]Center for Human Growth and Development, University of Michigan, Ann Arbor, MI, USA; [7]Institute for Social Research, University of Michigan, Ann Arbor, MI, USA; [8]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA and [9]Department of Psychology, Michigan State University, East Lansing, MI, USA

Abstract

Background. Callous-unemotional (CU) traits are critical to developmental, diagnostic, and clinical models of antisocial behaviors (AB). However, assessments of CU traits within large-scale longitudinal and neurobiologically focused investigations remain remarkably sparse. We sought to develop a brief measure of CU traits using items from widely administered instruments that could be linked to neuroimaging, genetic, and environmental data within already existing datasets and future studies.

Methods. Data came from a large and diverse sample ($n = 4525$) of youth (ages~9–11) taking part in the Adolescent Brain and Cognitive Development (ABCD) Study. Moderated nonlinear factor analysis was used to assess measurement invariance across sex, race, and age. We explored whether CU traits were distinct from other indicators of AB, investigated unique links with theoretically-relevant outcomes, and replicated findings in an independent sample.

Results. The brief CU traits measure demonstrated strong psychometric properties and evidence of measurement invariance across sex, race, and age. On average, boys endorsed higher levels of CU traits than girls and CU traits were related to, yet distinguishable from other indicators of AB. The CU traits construct also exhibited expected associations with theoretically important outcomes. Study findings were also replicated across an independent sample of youth.

Conclusions. In a large, multi-site study, a brief measure of CU traits can be measured distinctly from other dimensions of AB. This measure provides the scientific community with a method to assess CU traits in the ABCD sample, as well as in other studies that may benefit from a brief assessment of CU.

Childhood antisocial behavior (AB), including aggression, violence, and theft, puts youth at risk for chronic AB and criminality across the lifespan (Moffitt *et al.*, 2002). Moreover, the greater use of health and education services by children with AB, as well as the harmful effects of crime and AB on communities and families more broadly, confers a significant monetary cost to society (Rivenbark *et al.*, 2018). However, marked heterogeneity in patterns of youth AB has long made it difficult to identify those most at risk of exhibiting these behaviors into later development. More recently, findings across a growing body of literature suggest that Callous-Unemotional (CU) traits, which refer to a lack of empathic concern, shallow affect, and low moral regulation (Frick *et al.*, 2014*a*), can be used to delineate a subgroup of youth having a particularly recalcitrant and stable form of AB (for a review, see Frick *et al.*, 2014*a*, Waller *et al.*, 2017*b*). Indeed, CU traits were incorporated into the DSM-5 as a specifier for Conduct Disorder (i.e. with/without limited prosocial emotions; American Psychiatric Association, 2013), requiring clinicians to establish the presence or absence of CU traits among children in need of treatment for their clinically-significant AB. To date, a major impediment for this area of research is the limited number of large-scale, longitudinal investigations (Lahey, 2014; Raschle *et al.*, 2018) and neurobiologically focused studies (Frick, 2012; Frick *et al.*, 2014*b*) that include assessments of CU traits. Consequently, the use of innovative methods that would allow researchers to investigate CU traits by leveraging already existing data from high-quality studies, has the potential to offer enormous economic and scientific benefits (Curran and Hussong, 2009; Hussong *et al.*, 2013). The primary goal of the current study was to develop a brief measure of CU traits that could be used by researchers

to assess these features across a broad range of available datasets, as well as current and future studies in which a more comprehensive measure of CU may not be available.

The last decade has seen a wealth of psychometric work examining the reliability and construct validity of questionnaire measures of CU traits (Frick *et al.*, 2014*a*). Perhaps most prominently, a number of studies have developed and used brief five- to nine-item CU scales (Dadds *et al.*, 2005; Viding *et al.*, 2005; Willoughby *et al.*, 2014; Waller *et al.*, 2017*a*) to demonstrate that CU traits are distinguishable from oppositional and attention-deficit problems during early childhood (i.e. ages 3–7-years-old) (see Waller and Hyde, 2018 for a review). Notably however, similar work conducted during later childhood (i.e. ages 9–12) is comparably sparse. This is an important limitation of existing research, as late childhood represents a critical developmental window marked by pronounced changes in neurocognitive, emotional, and social functioning (Luna *et al.*, 2004; Choudhury *et al.*, 2006; Casey and Jones, 2010; Kouros *et al.*, 2010; Zalewski *et al.*, 2011); areas in which impairments are commonly linked to risk for the onset and maintenance of AB (Moffitt, 2018).

An additional and imperative question to address in the CU traits literature is whether measures of CU traits are comparable across youth. Indeed, establishing measurement equivalence is a critical priority across a number of scientific fields, as measurement inconsistencies have the potential to obscure study results and lead to spurious conclusions (Borsboom, 2006). More specifically, if scores on a measure of CU traits vary systematically as a function of some characteristic (e.g. sex, race, age), then we risk over- or under- classifying specific groups of youth as having CU traits. Moreover, any associations we find between CU traits and other meaningful variables of interest (e.g. neurocognitive functioning, intervention responsiveness) may not accurately reflect underlying differences in CU traits. Despite the need for studies of CU traits to explore potential measurement invariance across important demographic characteristics (i.e. sex, race, age) (Skeem and Cooke, 2010; Frick *et al.*, 2014*a*; Frick *et al.*, 2014*b*; Hawes *et al.*, 2018), remarkably few studies have actually established measurement equivalence of CU trait measures. Drawing on findings from research focused on assessing the construct of psychopathy more broadly, Tsang *et al.* (2014) found that items assessing psychopathic traits showed differential item functioning among Caucasian *v.* African American boys (Tsang *et al.*, 2014). This finding is consistent with studies that have explored psychopathy in adult samples, where individual items from psychopathy measures were shown to perform differently among African American *v.* Caucasian individuals (Cooke and Michie, 2001) and men *v.* women (Dotterer *et al.*, 2017). Other recent psychopathy focused research, however, has found little evidence of measurement inconsistiencies across race (Hawes *et al.*, 2018). As such, additional research is needed to explore whether measures of CU traits are commensurate across sex, race, and age in samples of youth.

Current study

The current study had four goals. The primary goal was to *develop a brief measure of CU traits* having strong practical application and generalizability. To do so, we used items from two of the most widely administered instruments available for assessing youth behaviors and mental health (De Clercq *et al.*, 2012; He *et al.*, 2013; Stone *et al.*, 2015; Oakland *et al.*, 2016): the Child Behavior Checklist (CBCL; Achenbach and Edelbrock, 1983) and the Strength and Difficulties Questionnaire (SDQ; Goodman, 1997). A series of descriptive and graphical approaches were used for initial item screening and moderated nonlinear factor analysis (MNLFA) (Curran *et al.*, 2014) was employed to explore whether items used to assess CU traits were commensurate across sex, race, and age. Unlike traditional approaches of assessing measurement equivalence, which can only test the influence of one covariate of interest at a time (e.g. sex, race, or age), MNLFA can produce person-specific CU trait scores when accounting for potential differences in both the latent factor and individual items as a function of all covariates of interest, simultaneously. Our second goal was to *test whether CU traits items form a separable factor from items that index other forms of childhood AB*, including conduct problems (CP), oppositional defiant disorder behaviors (ODD behaviors), and attention-deficit disorder behaviors (ADHD behaviors), consistent with prior work in the preschool period (Willoughby *et al.*, 2014; Waller *et al.*, 2015). Our third goal was to *examine convergent and divergent validity of our newly derived CU traits measure* by examining differential associations with external outcomes. In addition to associations with CP, ODD behaviors, and ADHD behaviors, we focused on two additional salient dimensions of behavior with strong theoretical relevance to CU traits: prosocial behaviors and anxiety (Frick *et al.*, 2014*a*). We hypothesized that CU traits would be uniquely related to higher levels of CP and ODD behaviors, and lower levels of prosocial behavior and anxiety. Our final goal was to *investigate the factor structure and convergent validity* of our newly derived measure relative to a frequently used measure of CU traits (the Inventory of CU Traits) among an independent sample of youth recruited via birth records.

We also note that data for this study comes from the landmark Adolescent Brain and Cognitive Development (ABCD) Study (https://abcdstudy.org/). The ABCD study represents the largest long-term investigation of brain development and child health in the USA; however, no measure of CU traits is administered as part of the standard ABCD protocol. Considering its wide-reach and anticipated impact, along with the ability to link neuroimaging, genetic, and environmental data, the ABCD study serves as an exemplar for the type of investigation in which assessments of CU are most needed, yet rarely available. An additional benefit of using data from the ABCD study for the development of a brief measure of CU is that the ABCD project is following an open science model. Thus, ABCD study data (including all code used in the current study) are being made publically available to the scientific research community, which has the potential to stimulate novel areas of investigation and would serve as an immense resource for researchers to explore questions around the etiology and developmental course of CU traits.

Methods

Participants

Participants in this study are enrolled in the ongoing longitudinal ABCD Study (https://abcdstudy.org/). Goals of the ABCD study include the recruitment of almost 12 000 healthy children, aged 9 to 10-years-old from across the USA, and follow-up into early adulthood. The ABCD study aims to transform our understanding of the genetic and environmental influences on brain development, structure and function, and their roles in substance use and other health outcomes (Volkow *et al.*, 2018). Participants across

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

the 21 study sites were recruited through elementary schools, both public (including charter schools) and private. The use of a multi-stage probability sampling approach (Heeringa *et al.*, 2017) was intended to ensure local randomization and representativeness within sites, while yielding a final combined ABCD sample that closely approximates national sociodemographics (Garavan *et al.*, 2018). The human research protection programs and institutional review boards at universities participating in the ABCD project approved all experimental and consenting procedures, and all participants or their legal guardian gave informed consent. The sample used for the present analyses included participants who completed their baseline assessment during the first year of the ABCD study and were included in the initial ABCD data release (ABCD Study Collection 2573; https://data-archive.nimh.nih.gov/abcd). The final sample for the present study consisted of 4524 youth, ages 9–11 ($M = 10.03$, s.d. = 0.60; 48% girls; 59% Caucasian, 10% African American, 20% Hispanic, 11% Other). Additional study information provided in Garavan *et al.* (2018).

The brief measure of CU traits derived in this study was also examined among an independent sample of youth taking part in the on-going Michigan Twin Neurogenetics Study (MTwiNS) which is part of the larger Twin Study of Behavioral and Emotional Development in Children (TBED-C; Burt and Klump, 2013), a large-scale study of twin families recruited via birth records and enriched to oversample families residing in disadvantaged neighborhoods. Participants from the MTwiNS sample thus far include 280 youth (140 twin pairs), ranging in age from 7 to 18-years-old (mean age = 13.45; S.D. = 2.69; 27.1% female; 70.7% Caucasian; for a detailed overview of the TBED-C study sample, see Burt and Klump, 2013).

Measures

Study Aim 1: derive a brief, reliable measure of CU traits

CU Traits. Items used to index CU traits were derived from parent-report versions of the CBCL (Achenbach and Edelbrock, 1983) and the SDQ (Goodman, 1997). The CBCL and SDQ are two of the commonly employed measures for assessing emotional and behavioral problems in youth, worldwide; each having been translated into more than 70 languages (De Clercq *et al.*, 2012; Jenkins *et al.*, 2014). Items from these measures were selected into an initial item pool based on their overlap with theoretical conceptualizations of CU traits and the item content of other measures used to assess CU traits in youth (Kimonis *et al.*, 2008; Waller *et al.*, 2015). This process resulted in the inclusion of four items from the CBCL (i.e. 'lack of guilt after misbehaving', 'teases others a lot', 'too fearful or anxious', 'feels too guilty') and three items from the prosociality subscale of the SDQ ('is considerate of others feelings', 'is helpful if someone is hurt or upset', 'offers to help others'). Items on each of these measures are rated on a 3-point scale ranging from 0 to 2. Two CBCL items ('too fearful or anxious', 'feels too guilty') and each of the SDQ items was reversed scored so that higher scores on all items were indicative of higher levels of CU traits. Due to each of these seven items exhibiting infrequent endorsement (<3%) of the third category option (i.e. a score of 2), item categories were collapsed into dichotomous scores (i.e. 0–1).

Study Aim 2: test whether CU traits items form a separable factor from items assessing CP, ODD behaviors, and ADHD behaviors

CP. We assessed CP via the 17-item DSM-Oriented 'Conduct Problem' scale from the CBCL. Items were rated on a 3-point scale ranging from 0 (*Not true*) to 2 (*Very true or Often true*) and summed such that higher scores represent increased levels of problems (e.g. 'breaks rules', 'steals', 'fights'; $\alpha = 0.77$; $\bar{x} = 1.01$, S.D. = 1.87). However, a single item (i.e. 'lack of guilt') was omitted from the CP scale in this study, in order to avoid overlap with item content used as part of the CU measure.

ODD Behaviors. We assessed ODD behaviors via the five-item DSM-Oriented 'Oppositional Defiant' scale from the CBCL. As before, items were rated on a 3-point scale ranging from 0 (*Not true*) to 2 (*Very true or Often true*) and summed such that higher scores represent increased levels of problems (e.g. 'argues', 'disobedient at home/school', and 'stubborn'; $\alpha = 0.80$; $\bar{x} = 1.75$, S.D. = 2.01).

ADHD Behaviors. We assessed ADHD behaviors via the seven-item DSM-Oriented 'Attention Deficit' scale of the CBCL. Items were rated on a 3-point scale ranging from 0 (*Not true*) to 2 (*Very true or Often true*) and summed such that higher scores represent increased levels of problems (e.g. 'impulsive', 'can't sit still', and 'can't concentrate'; $\alpha = 0.85$; $\bar{x} = 2.56$, S.D. = 2.92).

Study Aim 3: examine convergent and divergent validity of the CU traits measure

Prosociality. We assessed cross-informant associations between the overlapping CU traits and prosociality constructs by using a separate informant report (child self-report) of prosocial behaviors using the three items (i.e. 'I am considerate of others feelings', 'I am helpful if someone is hurt or upset', 'I offer to help others') from the youth-report prosocial scale of the SDQ (Goodman, 1997; Goodman and Goodman, 2009) that was assessed as part of the ABCD study ($\alpha = 0.60$).

Anxiety. Anxiety was assessed using the six-item DSM-Oriented 'Anxiety Problems Scale' of the CBCL (e.g. 'nervous', 'self-conscious', and 'fearful'; $\alpha = 0.62$).

Study Aim 4: independent sample replication

Inventory of CU Traits (ICU; Kimonis *et al.*, 2008). The parent report and youth report versions of the 24-item ICU scale were used to examine evidence of convergent validity of the brief CU traits measure derived in this study, among an independent sample of youth from MTwiNs (Burt and Klump, 2013). Consistent with prior factor analytic studies that provide support for a correlated three-factor model of the ICU (i.e. callousness subscale- 11 items; uncaring subscale- eight items; unemotional subscale- five items), item-level data was used to model latent ICU total and subscale scores for both, parent and youth report versions of the measure.

Analytic strategy

Study Aim 1: derive a brief, reliable measure of CU traits

We first employed traditional graphical and descriptive procedures to screen for poorly functioning CU traits items across the seven-item pool. Consistent with established guidelines, we explored item-total correlations (<0.30 considered to index poor discrimination), inter-item correlations (<0.20 considered to index items displaying low item homogeneity), and estimates of internal consistency (i.e. α of 0.30, 0.50, 0.80 used to establish low, moderate, and strong levels of internal consistency) (Clark and Watson, 1995). Exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) were then conducted to assess the unidimensionality of the CU traits construct. Subsequently, MNLFA was used to investigate differences on the CU traits factor

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

mean and variance, as well as DIF of item thresholds and factor loadings as a function of sex, race, and age, as well as their interaction effects (i.e. sex×race, sex×age, age×race). Only recently has MNLFA been introduced as a means for evaluating measurement invariance (Bauer, 2017); extending on traditional 'step-wise' procedures by allowing for simultaneous testing of whether item parameters (i.e. thresholds and factor loadings) remain invariant across multiple covariate influences.

To define the scale of the latent CU traits factor, the conditional mean and variance of the factor were set to 0 and 1 when all covariates were equal to zero (designating 10-year-old Caucasian males as the reference group). Item-level DIF was examined on an item-by-item basis, accounting for covariate effects on the CU traits factor mean and variance. Main effects of sex, race, and age on item thresholds and factor loadings were examined separately for each item, followed by analyses that included all two-way interactions between covariates (Curran *et al.*, 2014). To avoid inflating Type 1 error as a result of multiple comparisons while investigating DIF, we applied the Benjamini-Yekutieli correction (Benjamini and Yekutieli, 2001), with a false discovery rate of 5%. A final MNLFA model was established by retaining all significant covariate effects on the CU traits factor (mean and variance) and items (thresholds and loadings). Parameter estimates from this final model were used to produce *maximum a posteriori* (MAP) scale scores on the CU traits construct. Thus, CU traits MAP scores account for differences in the CU traits factor mean and variance, as well as item DIF, resulting from participant's sex, race, and age, along with any multiplicative effects. Unlike traditional summed score approaches, MAP scores provide unique information about individual differences by taking into account which items were endorsed and by whom, thus providing person-specific factor scores for the CU trait construct (Curran *et al.*, 2014).

Study Aim 2: test whether CU traits items form a separable factor from items assessing CP, ODD behaviors, and ADHD behaviors

We estimated and compared a series of five confirmatory factor analytic models to examine whether parent-reported CU traits could be distinguished from CP, ODD behaviors, and ADHD behaviors. First, items from all four measures were combined into a single-factor model, to test whether they represented a unidimensional construct, as opposed to distinct factors. Next, we estimated four two-factor models, which would imply that CU traits, CP, ODD behaviors, or ADHD behaviors could be distinguished as unique factors, but that the remaining items could not be differentiated. Finally, a four-factor model was estimated indicating that CU traits, CP, ODD behaviors, and ADHD behaviors could each be distinguished from each other. The DIFFTEST procedure in MPlus 7 (Muthén and Muthén, 1998–2012) was used to obtain corrected χ^2 difference tests for comparing these nested models.

Study Aim 3: examine convergent and divergent validity of the CU traits measure

CU traits MAP scores were used to evaluate convergent and discriminant validity of the CU traits measure by investigating associations with theoretically relevant constructs (e.g. CP, anxiety, child report prosociality). We explored bivariate associations between all study constructs. To establish the unique and cross-informant correlates of CU traits, we explored partial associations of CU traits with anxiety and youth self-report of prosociality, accounting for overlap with CP, ODD behaviors, and ADHD behaviors. Partial associations between CU and each of the three childhood externalizing behaviors constructs (i.e. CP, ODD behaviors, and ADHD behaviors) were examined by consecutively controlling for associations with the other two externalizing constructs (e.g. partial association between CU and CP accounts for effects of ODD and ADHD behaviors).

Study Aim 4: independent sample replication

Consistent with the approach outlined in '*Study Aim* 1', graphical and descriptive procedures were used to screen for poorly functioning CU traits items across the original seven-item pool using the TBED-C sample, while EFA and CFA models were conducted to assess the unidimensionality of the CU traits construct in this sample. To investigate evidence of convergent validity, a latent CU construct was created using the four items from the brief CU measure and bivariate correlations were examined with latent total and subscale scores of the parent and youth report ICU. For these analyses, the Mplus cluster option and Type = Complex was used to control for clustering of individuals within twin pairs.

All models in the current study were estimated using a mean- and variance-adjusted weighted least squares estimator appropriate for use with ordinal items (Flora and Curran, 2004) in Mplus 7 (Muthén and Muthén, 1998–2012). Mplus also allows for handling complex sampling and recruitment procedures, such as those implemented in the ABCD study (e.g. cluster correction for sibling pairs, stratification by study site). Model fit was assessed using the comparative fit index (CFI>0.90; Hu and Bentler, 1999) and root-mean-square error of approximation (RMSEA<0.05; Steiger, 1990).

Results

Study Aim 1: derive a brief, reliable measure of CU traits

A visual inspection of the scree plot produced by an exploratory factor analysis model provided strong evidence of a single dominant factor and this was also supported by the magnitude of the ratio (1.72) comparing the first ($\lambda = 2.36$) to second ($\lambda = 1.37$) eigenvalues. However, internal consistency for this original pool of seven items was not high ($\alpha = 0.63$) and examination of item-total correlations revealed three poorly discriminating items (<0.30; 'Teases others a lot', 'Too fearful or anxious', 'Feels too guilty'). Further inspection revealed that these three items also exhibited low inter-item correlations on average (<0.20) and two of these three items ('Teases others a lot', 'Feels too guilty') were endorsed by less than 10% of respondents. Thus, each of these items was consecutively removed from the total item pool. Subsequent investigation of the remaining items revealed satisfactory inter-item (>0.25) and item-total correlations (>0.30), acceptable model fit ($\chi^2(2) = 84.60$, $p < 0.001$, CFI = 0.99, TLI = 0.97, RMSEA = 0.09), and marked improvement in internal consistency ($\alpha = 0.73$;) (Table 1). Internal consistency among the four remaining items exceeded alpha values reported in several recent studies that have used a five-item CU traits scale (Willoughby *et al.*, 2011; Waller *et al.*, 2015).

Results obtained from the final MNLFA model are provided in Table 2. Findings revealed significant effects of participant's sex on the CU traits factor mean and variance, indicating that on average, boys endorsed higher levels of CU traits than girls. Accounting for these effects on the CU traits factor and after

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

Table 1. Item-level descriptive analyses

Item description	% Endorsed	Standardized loading (S.E.)	Item-total correlation	Average inter-item correlation
1. Not considerate of others feelings	24	0.89	0.64	0.50
2. Not helpful if someone is hurt or upset	17	0.93	0.62	0.48
3. Does not offer to help others	27	0.84	0.54	0.44
4. Lack of guilt after misbehaving	13	0.55	0.31	0.27

Table 2. Final MNLFA model results and covariate effects

		Covariate effects				
Reference parameter	Baseline	Sex	Age	African American	Hispanic	Other
CU						
Mean	0.00[a]	−0.58*** (0.08)	–	–	–	–
Variance	1.00[a]	0.22* (0.10)	–	–	–	–
Item 1. *Not considerate of others feelings*						
Threshold	2.88 (0.19)	–	–	–	–	–
Loading	3.64 (0.27)	–	–	–	–	−1.32*** (0.45)
Item 2. *Not helpful if someone is hurt or upset*						
Threshold	4.76 (0.36)	–	–	–	–	–
Loading	4.65 (0.40)	–	–	–	–	–
Item 3. *Does not offer to help others*						
Threshold	1.89 (0.09)	–	–	−1.52*** (0.36)	–	–
Loading	2.60 (0.14)	–	–	–	–	–
Item 4. *Lack of guilt after misbehaving*						
Threshold	2.33 (0.07)	–	–	0.67** (0.17)	0.52*** (0.14)	–
Loading	1.15 (0.07)	–	–	–	–	–

[a]Indicates parameter values are fixed to identify the model and set the scale of the latent variables.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

correcting for multiple comparisons, DIF analysis demonstrated a significant effect of race on the thresholds (i.e. *uniform* DIF) of two items ('does not offer to help others' and 'lack of guilt after misbehaving'), reflecting differences in the probability of endorsing these items at equivalent levels of underlying CU traits. That is, on average, lower levels of underlying CU traits were required for Caucasian youth, relative to African-American youth, to endorse the item 'does not offer to help others'. Alternatively, the opposite pattern was found for the threshold of the item 'lack of guilt after misbehaving', as African-American (and Hispanic) youth both demonstrated a higher probability of endorsing this item relative to Caucasian youth. Only a single instance of potential item loading (i.e. *non-uniform*) DIF was identified, between youth in the 'Caucasian' and 'Other' race categories for the item 'Not considerate of others feelings'. This suggests that the conditional dependency of endorsing this item may differ between youth in these two groups at different points across the CU trait continuum. Thus, to ensure comparability across different youth, these item thresholds and the single item loading which were identified as potential instances of DIF were allowed to differ across the specified race categories as noted above, in subsequent analyses.

Study Aim 2: test whether CU traits items form a separable factor from items assessing CP, ODD, and ADHD behaviors

A series of CFA models were compared to investigate whether parent report of CU traits could be appropriately distinguished from CP, ODD behaviors, and ADHD behaviors. Comparing these nested models, results from the DIFFTEST procedure indicated that the correlated four-factor model (i.e. separate CU, CP, ODD, and ADHD factors) demonstrated statistically superior fit to the single-factor model and two-factor models. The four-factor model provides support for the uniqueness of the CU traits construct, as well as distinguishing each of the constructs from each other. Information regarding model fit and model comparisons is summarized in Table 3. Consistent with previous findings, latent correlations between these factors were in the moderate-to-large range (r's = 0.31–0.68, $p < 0.001$), suggesting distinct but overlapping constructs (see Table 4).

Study Aim 3: examine convergent and divergent validity of the CU traits measure

MAP scores derived from the final MNLFA model, which accounted for covariate effects on the CU traits factor and item-

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

Table 3. Differentiation of CU traits factor from CP, ODD, and ADHD factors

Model	Description	CFI	RMSEA	Comparison	χ^2 (df)	*p*-value
1	1 factor	0.88	0.06	1 *v.* 6	2157.05 (6)	<0.001
2	2 factor (CU *v.* Other)	0.93	0.04	2 *v.* 6	1043.75 (5)	<0.001
3	2 factor (CP *v.* Other)	0.89	0.06	3 *v.* 6	1920.88 (5)	<0.001
4	2 factor (ODD *v.* Other)	0.89	0.06	4 *v.* 6	1841.71 (5)	<0.001
5	2 factor (ADHD *v.* Other)	0.92	0.05	5 *v.* 6	1166.21 (5)	<0.001
6	4 factor (CU, CP, ODD, ADHD)	0.96	0.03	–	–	–

CU, CU traits; CP, Conduct Problems; ODD, Oppositional Defiant Problems; ADHD, Attention Deficit Problems.

Table 4. Latent correlations between CU, CP, ODD and ADHD behaviors and theoretically relevant outcomes

	CU	CP	ODD	ADHD
CP	0.42*** (0.19***)			
ODD	0.43*** (0.18***)	0.68*** (0.47***)		
ADHD	0.31*** (0.02)	0.53*** (0.20***)	0.61*** (0.39***)	
Prosociality	−0.17*** (−0.15***)	−0.07*** (0.02)	−0.09*** (−0.01)	−0.08*** (−0.03*)
Anxiety	0.15*** (−0.08***)	0.36*** (0.05***)	0.45*** (0.21***)	0.45*** (0.23***)

CU, CU traits; CP, Conduct Problems; ODD, Oppositional Defiant Problems; ADHD, Attention Deficit Problems.
Notes: Partial correlations provided in parentheses. *<0.05, **<0.01, ***<0.001

level DIF, were used to examine the unique associations between CU traits and CP, ODD behaviors, and ADHD behaviors and other theoretically relevant outcomes (Table 4). First, as noted above, the bivariate associations between CU traits, CP, ODD and ADHD behaviors were moderate-to-large in magnitude. The correlation of CU traits with CP was somewhat attenuated after accounting for ODD and ADHD behaviors. Similarly, CU traits association with ODD behaviors was diminished, though still significant, when controlling for symptoms of CP and ADHD behaviors. Alternatively, the correlation between CU traits and ADHD behaviors became non-significant after accounting for symptoms of CP and ODD behaviors (Table 4). Second, CU traits, as well as CP, ODD and ADHD behaviors were each negatively related to youth self-report of prosociality. However, as hypothesized, only CU traits were uniquely related to low prosociality after accounting for overlap between CU traits and CP, ODD and ADHD behaviors. Finally, and in line with study hypotheses, there was evidence of a cross-over suppression effect for the association between CU traits and anxiety. That is, after controlling for symptoms of CP, ODD and ADHD behaviors, the positive association between CU traits and anxiety reversed directions, revealing a significant negative relationship between CU traits and anxiety (see Table 4).

Study Aim 4: independent sample replication

Finally, the brief CU measure derived in the current study was evaluated using data from the MTwiNS. Consistent with findings from the primary study analyses, the magnitude of the ratio (3.10) comparing the first ($\lambda = 3.07$) to second ($\lambda = 0.99$) eigenvalues provided support for a single dominant factor. In addition, the same three items as in the original analyses ('I tease others a lot', 'Too fearful or anxious', 'I feel too guilty') were identified as exhibiting low item-total (<0.30) and inter-item (<0.10) correlations. Removal of these items led to a noticeable improvement in internal consistency (i.e. change in coefficient α from $\alpha = 0.56$ to $\alpha = 0.69$) and evidence of a well-fitting model [$\chi^2(2) = 0.38$, $p = 0.82$, CFI = 1.00, TLI = 1.00, RMSEA = 0.00].

Finally, latent associations with the total and subscale scores on the parent and youth report ICU provided evidence supporting the convergent validity of the brief CU measure (see Table 5). These results showed the brief measure of CU traits to exhibit large correlations with the parent report ICU total ($r = 0.95$) and subscale scores (callousness $r = 0.92$; uncaring $r = 0.90$; unemotional $r = 0.48$). Moderate sized effects were also found for cross-informant associations with the youth report ICU (total $r = 0.36$; callousness $r = 0.30$; uncaring $r = 0.29$; unemotional $r = 0.22$). Notably, these effects are comparable to the magnitude of the associations between the parent and youth report versions of the ICU (see Table 5).

Discussion

A large and expanding body of literature has established that CU traits delineate a subgroup of youth with severe and recalcitrant forms of AB. In turn, CU traits have become an increasingly important diagnostic marker within clinical settings, highlighting the critical need for reliable measures of CU traits. However, assessments of CU in large-scale longitudinal investigations are notably lacking. The current study addressed this issue by developing a brief scale for assessing CU traits using commonly available measures, among youth taking part in the landmark ABCD study. This measure exhibited strong psychometric properties, independent sample replication, appropriate associations with theoretically relevant outcomes, and substantial overlap with measurement tools commonly used to assess CU traits in youth.

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

Table 5. Latent correlations between the brief CU traits measure and the ICU in an independent sample

	Brief CU measure	ICU parent report				ICU youth report		
	CU Traits	Total	Callous	Uncaring	Unemotional	Total	Callous	Uncaring
ICU parent report								
Total	0.95***							
Callous	0.92***	0.88***						
Uncaring	0.90***	0.92***	0.90***					
Unemotional	0.48***	0.71***	0.37***	0.50***				
ICU youth report								
Total	0.36***	0.31***	0.24***	0.33***	0.22***			
Callous	0.30***	0.22***	0.21***	0.25***	0.04	0.64***		
Uncaring	0.29***	0.28***	0.19***	0.29***	0.26***	0.85***	0.54***	
Unemotional	0.22***	0.17***	0.11	0.16*	0.25***	0.81***	0.66***	0.42***

Notes: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Development of a brief, reliable measure of CU traits

We used MNLFA to establish whether our measure was commensurate across youth based on sex, race, and age. This is critical because little existing CU traits research has focused on this topic and certainly not in such a large and diverse sample where these types of small/moderated effects can be well detected. We found that using our brief measure, boys exhibited higher levels of CU traits, on average, than girls. This parallels findings across a larger body of research that shows relatively large mean-level differences in CU traits across boys and girls (Cardinale and Marsh, 2017). However, despite this difference in *mean* levels of CU traits across sex, there was, overall, little evidence of DIF across the CU traits items examined in this study. That is, our measure of CU traits appeared to be relatively commensurate across boys and girls and youth of different ages and races.

We found only three instances of DIF, two of which represented instances of *uniform* DIF based on race. This type of DIF (as opposed to *non-uniform* DIF) suggests that an item's intercepts/thresholds function differently between groups of interest. That is, the probability of endorsing an item is greater for one group (e.g. Caucasians) relative to another group (e.g. African Americans, Hispanics) and this difference remains consistent (i.e. *uniform*) at all levels of the latent trait (e.g. low CU to high CU). Notably, findings were in opposite directions for the two different item thresholds that demonstrated *uniform* DIF in this study (i.e. probability of endorsing 'does not offer to help others' > among Caucasian youth; probability of endorsing 'lack of guilt' > among African American and Hispanic youth). One explanation for the different pattern of findings may be the wording of the two items (i.e. negatively- *v.* positively-worded), which a handful of studies have linked to different rates of endorsement of CU traits items (e.g. Ray *et al.*, 2016). It is less clear as to why the single instance of potential item loading (*non-uniform*) DIF may have occurred between youth in the 'Caucasian' and 'Other' categories, but not between any other comparisons of race groupings or for any other covariate, for this or any other item. Considering that analyses of race-based DIF of CU traits has not been previously explored between the race groups examined in this study, more research is needed before drawing specific conclusions on this issue. Nonetheless, it is important to note that in the current study, these instances of DIF were accounted for in the final MNLFA model by allowing these thresholds and loading to be unconstrained and vary across race when producing MAP scores.

Examining convergent and divergent validity of the CU traits measure

Consistent with extant literature, a correlated three-factor model was the best-fitting solution to our data showing that the CU traits items formed a distinct, albeit correlated factor, from items assessing childhood CP, ODD, and ADHD behaviors. Accordingly, items indexing a lack of concern for others feelings, acts of prosociality, and reduced guilt or remorse form a separable construct that is distinct from items that more generally index externalizing psychopathology and AB. Moreover, once the overlap between CU traits and these other constructs was taken into account, CU traits were uniquely related to lower child-reported prosociality, lower anxiety, and higher CP. Thus, our measure exhibited predictive validity in relation to risk for AB and construct validity in relation to theoretically-relevant constructs across informants. In particular, the parent report CU measure was uniquely related to child report of prosociality. Considering the overlapping item content, and centrality of prosocial behaviors to theoretical conceptualizations of CU (Frick *et al.*, 2014*a*), the cross-informant associations between these measures provide additional support for the construct validity of the CU traits measure derived in this study. Moreover, it is noteworthy that CU traits were related to lower anxiety after accounting for CP, ODD, and ADHD behaviors, consistent with theory that CU traits are indexed by a reduced arousal to cues of threat, including distress of others or stimuli that should provoke fear, including threat of punishment having broken the rules or transgressed against others (Blair, 2013).

Limitations

The current study was characterized by a number of strengths including the use of a large sample of youth closely approximating national sociodemographics, who were taking part in the ongoing

https://doi.org/10.1017/S0033291719000278 Published online by Cambridge University Press

ABCD study, as well as replication of findings within another well-sampled, independent study. However, study findings should also be considered within the context of several limitations. First, the CU traits measure was based solely on parent report data. Though the use of parent report to assess CU traits is common, future research should attempt to replicate these results across multiple informants. Second, given the circumscribed age range of study participants, it is unclear if these findings will generalize to preschool children or older adolescents. This restricted age range may have also reduced the likelihood of finding significant age-related moderation effects. In addition, although the brief measure of CU derived in this study exhibited relatively invariant properties across the participant categories that were examined, future research is needed that replicates these findings across a larger item pool that may more thoroughly tap into the underlying facets of CU traits. Finally, our study design was necessarily cross-sectional in nature. Thus, while we found expected associations between high scores on CU traits with indices of poor socioemotional functioning, future studies are needed to test the prospective predictive validity of our CU traits measure once subsequent data waves from the ABCD are available.

Conclusion

We derived a brief, four-item measure for CU traits that was commensurate across youth based on sex, race, and age, and exhibited evidence of construct (both discriminant and convergent) validity. Consistent with an open access framework, CU traits MAP scores (and accompanying code; see supplemental materials) will be made publicly available for any researchers to use and explore critical questions centered on the neuroetiology and environmental underpinnings of CU traits. These CU traits MAP scores provide unique information about individual differences by taking into account which items were endorsed and by whom, thus providing person-specific factor scores for CU traits and providing an exciting foundational platform for future research to explore CU traits.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291719000278.

Author ORCIDs. Samuel W. Hawes, 0000-0001-7173-1749

Acknowledgements. Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than I0,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA 041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147. A full list of supporters is available at https://abcdstudy.org/nih-collaborators. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/principal-investigators.html. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from the fast track data release. The raw data are available at http://dx.doi.org/10.15154/1412097. Instructions on how to create a NDA study are available at https ://dataarchive.nimh.nih.-gov/training/modules/study.html).

MTwiNS was supported by the National Institute of Mental Health and the Office of the Direction of the National Institute of Health via award UG3MH114249 to SAB and LWH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional funding for MTwiNS was provided by the Avielle Foundation via The Conway Family Award for Excellence in Neuroscience (to LWH and SAB) and institutional funding provided by the University of Michigan (to LWH). We thank the staff and familir their valuable contributions to the study. We are also greatly appreciative to Jordan Papp of the MTwiN study and Heidi Westerman of the MiND Lab for their data management efforts.

Conflict of interest. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Achenbach TM and Edelbrock C (1983) *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont.

Association AP (2013) *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Arlington, VA: American Psychiatric Association.

Bauer DJ (2017) A more general model for testing measurement invariance and differential item functioning. *Psychological Methods* **22**, 507.

Benjamini Y and Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29**, 1165–1188.

Blair RJR (2013) The neurobiology of psychopathic traits in youths. *Nature Reviews Neuroscience* **14**, 786.

Borsboom D (2006) When does measurement invariance matter? *Medical Care* **44**, S176–S181.

Burt SA and Klump KL (2013) The Michigan state university twin registry (MSUTR): an update. *Twin Research and Human Genetics* **16**, 344–350.

Cardinale EM and Marsh AA (2017) The reliability and validity of the Inventory of Callous Unemotional Traits: a meta-analytic review. *Assessment*. Advance online publication. doi: 10.1177/1073191117747392.

Casey B and Jones RM (2010) Neurobiology of the adolescent brain and behavior: implications for substance use disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* **49**, 1189–1201.

Choudhury S, Blakemore S-J and Charman T (2006) Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience* **1**, 165–174.

Clark LA and Watson D (1995) Constructing validity: basic issues in objective scale development. *Psychological Assessment* **7**, 309–319.

Cooke DJ and Michie C (2001) Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment* **13**, 171–188.

Curran PJ and Hussong AM (2009) Integrative data analysis: the simultaneous analysis of multiple data sets. *Psychological Methods* **14**, 81–100.

Curran PJ, Mcginley JS, Bauer DJ, Hussong AM, Burns A, Chassin L, Sher K and Zucker R (2014) A moderated nonlinear factor model for the development of commensurate measures in integrative data analysis. *Multivariate Behavioral Research* **49**, 214–231.

Dadds MR, Fraser J, Frost A and Hawes DJ (2005) Disentangling the underlying dimensions of psychopathy and conduct problems in childhood: a community study. *Journal of Consulting and Clinical Psychology* **73**, 400–410.

De Clercq B, Rettew D, Althoff RR and De Bolle M (2012) Childhood personality types: vulnerability and adaptation over time. *Journal of Child Psychology and Psychiatry* **53**, 716–722.

Dotterer HL, Waller R, Neumann CS, Shaw DS, Forbes EE, Hariri AR and Hyde LW (2017) Examining the factor structure of the Self-Report of Psychopathy Short-Form across four young adult samples. *Assessment* **24**, 1062–1079.

Flora DB and Curran PJ (2004) An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* **9**, 466–491.

Frick PJ (2012) Developmental pathways to conduct disorder: implications for future directions in research, assessment, and treatment. *Journal of Clinical Child and Adolescent Psychology* **41**, 378–389.

Frick PJ, Ray JV, Thornton LC and Kahn RE (2014*a*) 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**, 532–548.

Frick PJ, Ray JV, Thornton LC and Kahn RE (2014*b*) 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.

Garavan H, Bartsch H, Conway K, Decastro A, Goldstein R, Heeringa S, Jernigan T, Potter A, Thompson W and Zahs D (2018) Recruiting the ABCD sample: design considerations and procedures. *Developmental Cognitive Neuroscience* **32**, 16–22.

Goodman R (1997) The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry* **38**, 581–586.

Goodman A and Goodman R (2009) Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry* **48**, 400–403.

Hawes SW, Byrd AL, Kelley SE, Gonzalez R, Edens JF and Pardini DA (2018) Psychopathic features across development: assessing longitudinal invariance among caucasian and African American youths. *Journal of Research in Personality* **73**, 180–188.

He J-P, Burstein M, Schmitz A and Merikangas KR (2013) The strengths and difficulties questionnaire (SDQ): the factor structure and scale validation in US adolescents. *Journal of Abnormal Child Psychology* **41**, 583–595.

Heeringa SG, West BT and Berglund PA (2017) *Applied Survey Data Analysis*. New York: Chapman and Hall/CRC.

Hu L and Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* **6**, 1–55.

Hussong AM, Curran PJ and Bauer DJ (2013) Integrative data analysis in clinical psychology research. *Annual Review of Clinical Psychology* **9**, 61–89.

Jenkins LN, Demaray MK, Wren NS, Secord SM, Lyell KM, Magers AM, Setmeyer AJ, Rodelo C, Newcomb-Mcneal E and Tennant J (2014) A critical review of five commonly used social-emotional and behavioral screeners for elementary or secondary schools. *Contemporary School Psychology* **18**, 241–254.

Kimonis ER, Frick PJ, Skeem JL, Marsee MA, Cruise K, Muñoz LC, Aucoin KJ and Morris AS (2008) Assessing callous-unemotional traits in adolescent offenders: validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry* **31**, 241–252.

Kouros CD, Cummings EM and Davies PT (2010) Early trajectories of interparental conflict and externalizing problems as predictors of social competence in preadolescence. *Development and Psychopathology* **22**, 527–537.

Lahey BB (2014) What we need to know about callous-unemotional traits: Comment on Frick, Ray, Thornton, and Kahn (2014).

Luna B, Garver KE, Urban TA, Lazar NA and Sweeney JA (2004) Maturation of cognitive processes from late childhood to adulthood. *Child Development* **75**, 1357–1372.

Moffitt TE (2018) Male antisocial behaviour in adolescence and beyond. *Nature Human Behaviour* **2**, 177–186.

Moffitt TE, Caspi A, Harrington H and Milne BJ (2002) Males on the life-course-persistent and adolescence-limited antisocial pathways: follow-up at age 26 years. *Development and Psychopathology* **14**, 179–207.

Muthén LK and Muthén BO (1998–2012) *Mplus User's Guide*, 7th Edn. Los Angeles, CA: Muthén & Muthén.

Oakland T, Douglas S and Kane H (2016) Top ten standardized tests used internationally with children and youth by school psychologists in 64 countries: a 24-year follow-up study. *Journal of Psychoeducational Assessment* **34**, 166–176.

Raschle NM, Menks WM, Fehlbaum LV, Steppan M, Smaragdi A, Gonzalez-Madruga K, Rogers J, Clanton R, Kohls G and Martinelli A (2018) Callous-unemotional traits and brain structure: sex-specific effects in anterior insula of typically-developing youths. *NeuroImage: Clinical* **17**, 856–864.

Ray JV, Frick PJ, Thornton LC, Steinberg L and Cauffman E (2016) Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological Assessment* **28**, 394.

Rivenbark JG, Odgers CL, Caspi A, Harrington H, Hogan S, Houts RM, Poulton R and Moffitt TE (2018) The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *Journal of Child Psychology and Psychiatry* **59**, 703–710.

Skeem JL and Cooke DJ (2010) Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment* **22**, 433–445.

Steiger JH (1990) Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research* **25**, 173–180.

Stone LL, Janssens JM, Vermulst AA, Van Der Maten M, Engels RC and Otten R (2015) The Strengths and Difficulties Questionnaire: psychometric properties of the parent and teacher version in children aged 4–7. *BMC Psychology* **3**, 4.

Tsang S, Piquero AR and Cauffman E (2014) An examination of the Psychopathy Checklist: Youth Version (PCL: YV) among male adolescent offenders: an item response theory analysis. *Psychological Assessment* **26**, 1333.

Viding E, Blair RJR, Moffitt TE and Plomin R (2005) Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry* **46**, 592–597.

Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, Pérez-Stable EJ, Riley WT, Bloch MH and Conway K (2018) The conception of the ABCD study: from substance use to a broad NIH collaboration. *Developmental Cognitive Neuroscience* **32**, 4–7.

Waller R and Hyde LW (2018) Callous-unemotional behaviors in early childhood: the development of empathy and prosociality gone awry. *Current Opinion in Psychology* **20**, 11–16.

Waller R, Hyde LW, Grabell AS, Alves ML and Olson SL (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.

Waller R, Shaw DS and Hyde LW (2017*a*) Observed fearlessness and positive parenting interact to predict childhood callous-unemotional behaviors among low-income boys. *Journal of Child Psychology and Psychiatry* **58**, 282–291.

Waller R, Shaw DS, Neiderhiser JM, Ganiban JM, Natsuaki MN, Reiss D, Trentacosta CJ, Leve LD and Hyde LW (2017*b*) Towards an understanding of the role of the environment in the development of early callous behavior. *Journal of Personality* **85**, 90–103.

Willoughby MT, Waschbusch DA, Moore GA and Propper CB (2011) Using the ASEBA to screen for callous unemotional traits in early childhood: factor structure, temporal stability, and utility. *Journal of Psychopathology and Behavioral Assessment* **33**, 19–30.

Willoughby MT, Mills-Koonce RW, Gottfredson NC and Wagner NJ (2014) Measuring callous unemotional behaviors in early childhood: factor structure and the prediction of stable aggression in middle childhood. *Journal of Psychopathology and Behavioral Assessment* **36**, 30–42.

Zalewski M, Lengua LJ, Wilson AC, Trancik A and Bazinet A (2011) Emotion regulation profiles, temperament, and adjustment problems in preadolescents. *Child Development* **82**, 951–966.