

Shared and Unique Variances of Interpersonal Callousness and Low Prosocial Behavior

Alan J. Meehan
King's College London

David J. Hawes
University of Sydney

Randall T. Salekin
University of Alabama

Edward D. Barker
King's College London

Although low prosocial behavior (LPB) items have been incorporated into youth measures of callousness, it remains unclear from current factor analytic findings whether callous traits and LPB are best operationalized as a common construct, or distinct dimensions. Using data from a population-representative birth cohort ($N = 5,463$), this study compared 4 latent factor structures for interpersonal callousness (IC; 6 items) and LPB (5 items) at age 13: (a) unidimensional; (b) two-factor; (c) higher-order (with 2 subfactors); and (d) bifactor (1 general and 2 specific residual factors). Alternative models distinguishing positively and negatively worded items were tested for comparative purposes. To assess the external validity of the factors that emerged from the best-fitting model, associations with early parenting styles and psychiatric comorbidities were examined. A bifactor model, achieving invariance for males and females, offered the best fit for these data. However, additional bifactor-specific indices suggested that the specific IC factor did not offer a unique contribution to the total variance over and above the general factor (IC/LPB). Of the remaining factors, IC/LPB was associated with higher levels of harsh parenting, externalizing and internalizing disorder, and social-cognitive difficulties, and lower levels of warm parenting. The LPB factor, meanwhile, was associated with greater social-cognitive difficulties and externalizing disorder, and lower maternal warmth, evoking a phenotype that may be more indicative of the autism spectrum than IC. These findings suggest that the shared variance underlying IC and LPB taps a severe psychiatric phenotype, while the residual variance for LPB may represent a distinct profile of social-cognitive dysfunction.

Public Significance Statement

Using factor analysis, we identified substantial shared variance underlying items measuring interpersonal-callous traits and low prosocial behavior in early adolescence. This was represented by a general factor, characterized by poorer parenting and greater comorbid child psychopathology. Given the apparent overlap between these two constructs, incorporating criteria to assess prosocial expression, or a lack thereof, may further enhance diagnostic measures of callousness in relation to childhood conduct problems.

Keywords: interpersonal callousness, low prosocial behavior, child psychopathy, psychopathology, Avon Longitudinal Study of Parents and Children (ALSPAC)

Supplemental materials: <http://dx.doi.org/10.1037/pas0000675.supp>

This article was published Online First December 10, 2018.

Alan J. Meehan, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London; David J. Hawes, School of Psychology, University of Sydney; Randall T. Salekin, Department of Psychology, University of Alabama; Edward D. Barker, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London.

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The United Kingdom Medical Research Council and Wellcome Trust (102215/2/13/2) and University of Bristol provide core support for

ALSPAC. Alan J. Meehan is supported by the Economic and Social Research Council (ES/J500057/1) and the National University of Ireland. Research reported in this publication was supported by awards from the National Institute of Child Health and Human Development (R01HD068437) and Economic and Social Research Council (ES/R005516/1) to Edward D. Barker.

Correspondence concerning this article should be addressed to Alan J. Meehan, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AF, or to Edward D. Barker, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AF. E-mail: alan.meehan@kcl.ac.uk or ted.barker@kcl.ac.uk

The most significant change to the diagnostic criteria for conduct disorder (CD) in *DSM-5* was the addition of a “with limited prosocial emotions” specifier, based on the presence of the following features: lack of remorse or guilt, callous-lack of empathy, unconcern about performance, and shallow or deficient affect (American Psychiatric Association, 2013). Although the decision to use only affective characteristics to specify CD has been criticized as potentially being too narrow (Salekin, 2016a, 2016b, 2017), in general, the case for incorporating even some psychopathic traits into diagnostic criteria has merit. This is because childhood studies of psychopathic traits have shown that youth who exhibit these traits present a more severe and stable trajectory of conduct problems (CP) and aggressive behavior, and show distinct risk factors compared to CP youth low on psychopathic traits (see Frick, Ray, Thornton, & Kahn, 2014, for a comprehensive review). Although this specifier dictates that individuals must first be diagnosed with CD, youth with elevated psychopathic traits in the absence of co-occurring conduct problems also exhibit higher levels of externalizing and, in some cases, internalizing problems compared with youth low on psychopathic traits and CD (Pardini, Stepp, Hipwell, Stouthamer-Loeber, & Loeber, 2012; Rowe et al., 2010).

The fact that *DSM-5*’s specifier for psychopathic symptoms is termed “with limited prosocial emotions” reflects a commonly held assumption that psychopathic traits are reflective of an interpersonally callous style, which includes a self-centered lack of concern or consideration for others (i.e., low prosocial behavior). Prosocial behavior here refers to voluntary action intended to benefit another person, including helping, sharing, and comforting behaviors (Eisenberg, Fabes, & Spinrad, 2006). Such behaviors are cross-sectionally and prospectively linked with better social adjustment and educational attainment (e.g., Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; X. Chen et al., 2002; Gerbino et al., 2018; Vitaro, Brendgen, Larose, & Tremblay, 2005). On the other hand, low prosocial behavior, as with psychopathic traits, is associated with more aggressive and delinquent behaviors (Carlo et al., 2014; Eivers, Brendgen, Vitaro, & Borge, 2012; Kokko & Pulkkinen, 2000). Indeed, longitudinal analyses of prosociality across childhood and adolescence consistently identify “low prosocial” developmental trajectories, membership of which is associated with higher levels of aggression and externalizing behavior compared to medium- or high-prosocial trajectory groups (Flynn, Ehrenreich, Beron, & Underwood, 2015; Kokko, Tremblay, Lacourse, Nagin, & Vitaro, 2006; Nantel-Vivier, Pihl, Côté, & Tremblay, 2014). However, longitudinal findings pertaining to internalizing symptoms are less clear (Nantel-Vivier et al., 2014). Elsewhere, children with significant social-cognitive deficits or formal diagnoses of autism spectrum disorder (ASD) generally display less prosocial behavior than their typically developing counterparts, in both cohort (A. Goodman, Lamping, & Ploubidis, 2010; Russell et al., 2012) and experimental designs (Jameel, Vyas, Bellesi, Roberts, & Channon, 2014). In contrast, while both psychopathic and autistic traits are superficially characterized by empathic difficulty, correlations between measures of the two constructs are generally modest, while one childhood twin study found little “phenotypic overlap” between autistic social and communication deficits and psychopathic traits in relation to relative genetic and environmental influences (O’Nions et al., 2015).

In terms of etiological influences, meanwhile, psychopathic traits and low prosocial behavior show similar profiles of early environmental exposure, particularly early parenting experiences including, for instance, harsh and inconsistent discipline (Hastings, Utendale, & Sullivan, 2007; Knafo & Plomin, 2006; Pardini, Lochman, & Powell, 2007; Waller et al., 2012) and a lack of parental warmth (Carlo, Mestre, Samper, Tur, & Armenta, 2010; Day & Padilla-Walker, 2009; Eisenberg, Spinrad, & Knafo-Noam, 2015; Kroneman, Hipwell, Loeber, Koot, & Pardini, 2011; Pasalich, Dadds, Hawes, & Brennan, 2011). However, the relation between parenting and psychopathic traits can be somewhat complex (Edens, Skopp, & Cahill, 2008; Hipwell et al., 2007; Oxford, Cavell, & Hughes, 2003).

The potential association between psychopathic traits and low prosocial behavior more broadly, as suggested by their similarity in terms of psychiatric comorbidity and early risk exposure, is reflected in formal measures of these constructs. It should be noted here that childhood studies of psychopathic traits have traditionally focused on Factor 1 of the two-factor model of psychopathy originally identified in the Psychopathy Checklist (PCL; Harpur, Hare, & Hakstian, 1989) and its revision (PCL-R; Hare, 2003). This comprises interpersonal (e.g., superficial charm, deceitful behavior) and affective features (e.g., lack of empathy/guilt, or “callous-unemotional” traits). Although more recent three- and four-factor models separate these facets, the interpersonal-affective dimension is generally referred to as interpersonal callousness (IC) in the child literature, and hereafter (e.g., Byrd, Hawes, Loeber, & Pardini, 2018; Byrd, Loeber, & Pardini, 2012; Pardini, Obradovic, & Loeber, 2006).

Cross-sectional and longitudinal childhood studies consistently find associations between higher levels of callousness and lower levels of prosocial behavior; these are generally modest, however, implying some degree of nonshared variance (Barker, Oliver, Viding, Salekin, & Maughan, 2011; Meehan, Maughan, Cecil, & Barker, 2017; Moran et al., 2009). At an item level, common prosocial behavior measures, such as the prosocial subscale of the Strengths and Difficulties Questionnaire (SDQ; R. Goodman, 1997), generally only assess the frequency of observed prosocial action in a child’s everyday activities; for example, whether they are “helpful if someone is hurt, upset, or ill” or “kind to younger children.” In contrast, measures of a callous interpersonal style (i.e., IC), such as the Antisocial Process Screening Device (APSD; Frick & Hare, 2001), focus more on the child’s affective profile (e.g., “not concerned about others’ feelings;” “hides feelings and emotions from others”). In particular, these scales are more concerned with the respondent’s emotional response to their behavior, or lack thereof (e.g., “doesn’t feel bad or guilty when does something wrong”). In this way, although a subgroup of those with poor prosocial functioning may also have high levels of IC, with distinct affective features underpinning their actions, not all “low prosocial” youth will necessarily display IC. Indeed, there may be other etiological influences on the broad “low prosocial” construct; for example, social functioning can be impaired as a result of cognitive deficits, such as an inability to infer the mental states of others, as seen in relation to ASD (Jameel et al., 2014; Pasalich, Dadds, & Hawes, 2014).

Notwithstanding these potential unique features, childhood studies have identified significant overlap between callous and

“low prosocial” measures. Specifically, two parallel studies of callousness by independent research groups (Dadds, Fraser, Frost, & Hawes, 2005; Viding, Blair, Moffitt, & Plomin, 2005) created scales that each incorporated “callous” items from the APSD (e.g., “no guilt,” “does not show feelings or emotions,” “breaks promises”) and reverse-scored items from the SDQ’s prosocial scale (e.g., “unhelpful if someone is hurt, upset, or ill,” “not kind to younger children”). In particular, Dadds et al. (2005) reported that the *de novo* factor combining these items, based on principal components analysis, showed higher reliability than the APSD’s original Callous-Unemotional subscale. In addition, these (low) prosocial items did not significantly load on an Antisocial factor, which captured more severe behavioral problems (e.g., lying, fighting, stealing), suggesting that the propensity to be uncaring to others was more indicative of a callous interpersonal style (i.e., IC, or Factor 1 within the two-factor model for psychopathy). These combined “callous-low prosocial” scales have since been used extensively in studies of youth callousness (e.g., Dadds et al., 2006; Fontaine, McCrory, Boivin, Moffitt, & Viding, 2011; Hawes, Price, & Dadds, 2014; Pasalich et al., 2011).

While the reliability of such a measure may support the notion that a lack of prosocial behavior is central to IC, factor analyses of a more comprehensive measure of callousness, the Inventory of Callous-Unemotional Traits (ICU; Frick, 2004), have identified two-factor models, comprised of a “callous” factor alongside an “uncaring” or “(low) empathic-prosocial” factor, which is itself made up of reverse-scored items originally intended to capture prosociality (e.g., “apologizes to people s/he has hurt,” “does things to make others feel good,” “tries not to hurt others’ feelings,” S. W. Hawes et al., 2014; Waller et al., 2015; Willoughby, Mills-Koonce, Waschbusch, & Gottfredson, 2015). However, item response analysis has suggested that this two-factor solution may be an unintended byproduct of the fact that responses to “uncaring” items were all reversed, unlike “callous” items, which were exclusively positively worded (Ray, Frick, Thornton, Steinberg, & Cauffman, 2016). Consequently, the separation of these factors may reflect method variance related to differences in the underlying response or endorsement patterns for these differently worded items, rather than a meaningful conceptual distinction.

In summary, childhood interpersonal-callous traits and low prosocial behavior have broadly similar profiles of comorbidity and etiology, and have previously shown construct validity and clinical utility when combined in a single measure. However, the extent to which “interpersonally callous” and “low prosocial” dimensions represent distinct constructs, or a common underlying factor, remains unclear. Therefore, this study had three main aims. First, we sought to parse the relative shared and unique variances underlying IC and low prosocial behavior (termed LPB hereafter), by testing competing factor structures from previous studies of callousness. Specifically, we compared (a) unidimensional, (b) two-factor, (c) higher-order, and (d) bifactor models. Given the previous utility of combined callous-low prosocial measures (Dadds et al., 2005; Viding et al., 2005), and the dominance of the two-factor model for psychopathy to which these two constructs may correspond (Hare, 2003), we hypothesized that substantial shared variance between IC and LPB would be identified. Second, given evidence from previous youth studies that boys score higher on psychopathic traits, and lower on prosociality, than girls (e.g., Eisenberg et al., 2006; Essau, Sasagawa, & Frick, 2006; Viding,

Simmonds, Petrides, & Frederickson, 2009), we also evaluated the extent to which sex differences impacted on our own measures of these constructs, and on our best-fitting latent factor model. Finally, we sought to better characterize our best-fitting IC and LPB factors by examining profiles of association with common correlates of both IC and LPB. Specifically, we compared the resulting latent factors on external measures of harsh and warm parenting in early childhood, and psychiatric comorbidity (i.e., externalizing problems, internalizing problems, and social-cognitive difficulties) throughout childhood.

Method

Participants

Data were drawn from the Avon Longitudinal Study of Parents and Children (ALSPAC), a population-representative British birth cohort established to understand how genetic and environmental characteristics influence health and development in parents and children (Boyd et al., 2013). Pregnant women resident in the former Avon Health Authority with expected delivery dates between April 1, 1991 and December 31, 1992 were eligible for recruitment. This yielded 14,541 pregnancies, of which 13,988 singletons/twins were alive at 12 months of age. When compared with 1991 National Census data, this cohort has been found to be broadly representative of both the Avon catchment area and the wider British population (Fraser et al., 2013). Ethical approval was obtained from the ALSPAC Law and Ethics Committee and various Local Research Committees. The study website contains details of all available data, through a fully searchable data dictionary: <http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>.

From the original ALSPAC cohort, 5,463 participants (49.8% female) had complete data for interpersonal callousness and low prosocial behavior at age 13 years and were selected for analysis. Although evenly distributed in terms of sex, it should be noted that the ethnic composition of this analytic sample was 98.6% White; however, this is broadly consistent with the Avon region at the time of recruitment, as well as that of the initial enrolled sample (96.1% White; Boyd et al., 2013). With regard to socioeconomic status (SES), attrition within ALSPAC over time has generally resulted in a loss of younger, more socially disadvantaged mothers at follow-up. For example, 9% of mothers in the current sample were classified as “low SES,” based on Classes IV and V of the Registrar General’s social class scale (Office of Population Censuses & Surveys, 1991), compared with 12% of the initial sample. Multivariate logistic regression and odds ratios (*ORs*) were used to examine whether low SES (*OR* = 1.40, 95% CI [1.23, 1.60]), early parenthood (19 years or younger; *OR* = 2.89, 95% CI [2.15, 3.89]), and low maternal educational attainment (basic school-leaving/vocational qualifications only; *OR* = 1.82, 95% CI [1.64, 2.01]) predicted exclusion from our analytic sample. All three variables were significantly associated with exclusion. However, it should be noted that a previous study of attrition bias in ALSPAC found that although attrition impacted the prevalence of psychiatric disorders, associations between risks and outcomes remained intact, and were likely to be conservative of the true population effects (Wolke et al., 2009).

Measures

Interpersonal callousness (IC). A six-item measure was completed by mothers when their child was 13-years-old (Moran, Ford, Butler, & Goodman, 2008). On a 5-point scale (0 = *not at all* to 4 = *always*), items rated how much the child: (a) makes a good impression at first, which people tend to see through after getting to know him/her; (b) has shallow or fast-changing emotions; (c) is usually genuinely sorry if s/he has hurt someone or acted badly (reversed); (d) can seem cold-blooded or callous; (e) keeps promises (reversed); and (f) is genuine in his or her expression of emotions (reversed). Initial item selection was informed by previous factor analyses of scales measuring traits from Factor 1 of the PCL-R (i.e., interpersonal and affective characteristics), which is currently the international standard for the assessment of psychopathy and has played a dominant role in the establishment of childhood measures (Frick, Bodin, & Barry, 2000; Frick, O'Brien, Wootton, & McBurnett, 1994; Hare, 2003). Validating the present scale in a sample of 182 clinic-referred or school-recruited children who scored highly for externalizing disorders, Moran et al. (2009) reported a high correlation ($r = .81$) with the APSD's (Frick & Hare, 2001) Callous Unemotional subscale. Internal consistency was acceptable within the current sample ($\alpha = .75$).

Low prosocial behavior (LPB). The prosocial subscale of the SDQ (R. Goodman, 1997), previously incorporated into childhood studies of callous traits (Dadds et al., 2005; Viding et al., 2005), was completed by mothers when the child was aged 13 years. Five items assessed the following behaviors on a 3-point scale (0 = *not true* to 2 = *certainly true*): (a) considerate of other's feelings; (b) shares readily with other children; (c) helpful if someone is hurt, upset, or ill; (d) kind to younger children; and (e) volunteers to help others. Responses were reversed, whereby higher scores captured lower prosociality. Internal consistency was acceptable ($\alpha = .71$).

Early parenting. Harsh parenting was assessed by two items each at ages 2 and 4, asking the mother "When at home with your child, how often do you": (a) shout at him/her?; and (b) slap him/her? (1 = *rarely/never* to 5 = *every day*). Scores for both ages were combined into a single latent factor using confirmatory factor analysis (CFA; $\alpha = .71$).

Warm parenting at age 2 was assessed by maternal ratings of the extent to which they (a) really love the toddler; (b) have pleasure in watching the child grow; and (c) feel the child provides great joy (1 = *feel never* to 4 = *feel exactly*). At age 4, maternal warmth was captured by five items, asking how much the mother (a) sings to; (b) reads to; (c) plays with toys with; (d) plays imitation games with; and (e) engages in physical play with their child (1 = *never* to 5 = *nearly every day*). As with harshness, scores at ages 2 and 4 were combined into one latent variable ($\alpha = .69$).

Psychiatric comorbidities. Externalizing and internalizing disorders at 7, 10, and 13 years were drawn from the Development and Well Being Assessment (DAWBA), originally developed for the British Child Mental Health Surveys (R. Goodman, Ford, Richards, Gatward, & Meltzer, 2000). Preliminary DSM-IV psychiatric diagnoses were generated from parent-reported symptoms using a well-defined computerized algorithm (see <http://www.dawba.com>), producing six-level ordered-categorical "probability bands" for each disorder, ranging from <0.1% to >70% probability of diagnosis. These bands functioned well as ordered-categorical

measures when evaluated in two large-scale national samples, showing dose-response associations with mental health service contacts, and similar associations with potential risk factors as clinician-rated diagnoses (A. Goodman, Heiervang, Collishaw, & Goodman, 2011). For externalizing disorder, we incorporated diagnoses for attention-deficit/hyperactivity disorder (ADHD), conduct disorder (CD), and oppositional defiant disorder (ODD) using a latent factor structure, while internalizing disorder was represented by a latent variable made up of anxiety and depression (see Supplemental Figure S1).

Social-cognitive difficulties at ages 7, 10, and 13 were captured by parent ratings on the 12-item Social Communication Disorders Checklist (SCDC; Skuse, Mandy, & Scourfield, 2005). This measured social reciprocity and communication deficits over the past 6 months, assessing the verbal and nonverbal social traits characterizing autism spectrum disorder (ASD). Items were scored from 0 (*not true*) to 2 (*very/often true*); thus, higher scores implied greater social-cognitive dysfunction. This scale shows high sensitivity and specificity for ASD diagnosis using a score of ≥ 9 out of 24 (Skuse et al., 2009), and showed good internal consistency at all time-points ($\alpha = .87-.89$).

Statistical Analyses

Step 1: Model fitting. Using Mplus v7.11 (Muthén & Muthén, 2012), we examined the best-fitting CFA structure for IC (6 items) and LPB (5 items). Four types of models were tested: (a) unidimensional (i.e., one-factor); (b) two (correlated) factors; (c) higher-order (with two first-order subscale factors); and (d) bifactor (one general and two specific factors, with covariances between all factors fixed to zero; Brown, 2006). Additionally, given recent findings highlighting the influence of item wording on previous factor analyses of the callousness construct (Ray et al., 2016), we also compared these models with an alternative two-factor solution, which specified two factors simply based on the direction in which items were worded (i.e., positive vs. negative wording), as well as a further three-factor solution, separating LPB items, positively worded IC items, and negatively worded IC items. We used mean- and variance-adjusted weighted least squares estimation (WLSMV) for all models, as recommended for ordinal data with less than five response categories, as with the 3-point LPB items (Rhemtulla, Brosseau-Liard, & Savalei, 2012). The χ^2 goodness-of-fit statistic was used to assess absolute model fit, where a nonsignificant value (i.e., $p > .05$) indicated good model fit. Relative fit was assessed using the comparative fit index (CFI) and Tucker-Lewis index (TLI), with values $\geq .95$ indicating good fit for both (Hu & Bentler, 1999). We also used the root mean square error of approximation (RMSEA), where values $< .05$ indicated good fit, and values $< .08$ demonstrated acceptable fit (Browne & Cudeck, 1993). Where models were nested, comparisons in model fit were evaluated using a chi-square difference test, via the DIFFTEST Mplus command.

In relation to the bifactor solution, concerns have recently been raised about the use of these models to represent a latent underlying dimension, alongside orthogonal specific factors. In traditional bifactor models, the specific factors represent the residual correlations that remain once a general factor has been extracted (Holzinger & Swineford, 1937). However, when applied to measures of psychopathology, the extent to which these residual fac-

tors represent meaningful constructs, or simply capture outstanding measurement error or noise, often remains untested (Bonifay, Lane, & Reise, 2017). Moreover, the inclusion of this nuisance variance may lead to “overfitting” of the data, which could account for the superior fit that is often reported for these models using conventional fit statistics (Reise, Kim, Mansolf, & Widaman, 2016). Consequently, additional indices have been proposed to evaluate the appropriateness of a bifactor structure for the data, based on the variance that is explained by general or specific factors.

We estimated several indices for the bifactor model using the Omega program (Watkins, 2013); full details of their calculation and interpretation can be found in Rodriguez, Reise, and Haviland (2016a, 2016b). First, reliability estimates were provided by several variants of the omega coefficient. Specifically, omega (ω), viewed as a latent variable analogue to coefficient alpha, represents the proportion of variance in the unit-weighted total score attributable to *all modeled sources* of common variance (i.e., the general and all group factors). Similarly, omega subscale (ω_s) estimates the proportion of variance in the composite score for a given subscale attributable to all sources of common variance (i.e., the general and particular group factor). The omega hierarchical (ω_H) coefficient, in contrast, measures the proportion of systematic variance in the total score that can be specifically attributed to individual differences on the *general* factor. Similarly, omega hierarchical subscale (ω_{HS}) assesses the proportion of variance in a subscale score accounted for by its *intended specific* factor, after controlling for the effects of the general factor. Thus, the discrepancy between ω and ω_{HS} values for a specific factor illustrates the scale to which the common variance explained by that subscale is accounted for by variance in the general factor. There are no absolute standards for ω_H or ω_{HS} : although values $>.75$ are preferred, values $>.50$ can still indicate that a factor provides sufficiently reliable unique variance (Reise, Bonifay, & Haviland, 2013).

Second, construct reliability, or the extent to which a latent factor is well-defined by its underlying items, was measured using H . A high H value (i.e., $>.70$) for a latent variable means that it is more likely to be reliably replicated across samples. Third, explained common variance (ECV) and percentage of uncontaminated correlations (PUC) evaluated the extent to which the bifactor model was genuinely multidimensional. More specifically, ECV is the proportion of the variance explained by all factors that is explained by the general factor, while PUC indicates the proportion of correlations between items that are influenced by the general factor. Where both ECV and PUC are $>.70$, this indicates a strong general factor, such that estimating a unidimensional model may offer a more parsimonious solution than the bifactor structure, without introducing any unwanted bias in the factor loadings (Rodriguez et al., 2016a). Finally, the potential biasing effect from forcing multidimensional data to be unidimensional was evaluated using the relative parameter bias, or the difference between an item's factor loading in the unidimensional model and its loading on the general factor in a bifactor model, divided by its general factor loading in the bifactor. A parameter bias of $<10\%$ – 15% between loadings suggests little meaningful bias between unidimensional and bifactor models.

Step 2: Sex differences. Differences between male and female participants were assessed on two levels. First, we tested

mean differences between males and females on derived total scores for IC and LPB using independent-sample t tests. Second, we tested multiple-group measurement invariance, or the equivalency of our best-fitting model across the sexes. This determines whether the latent factor structure of a set of items is understood similarly for males and females alike, which is seen as a prerequisite for making meaningful group comparisons. Based on the procedure outlined by Brown (2006), we sequentially tested three levels of measurement invariance, with increasing equality constraints introduced across the groups at each successive level. As a first step, CFA models were estimated separately for males and females. We then tested configural invariance, that is, whether the overall latent factor structure was the same across groups, by estimating the unconstrained model simultaneously for males and females. As in Step 1, goodness-of-fit was denoted by χ^2 , CFI, TLI, and RMSEA statistics. Next, we tested metric (or “weak”) invariance by constraining factor loadings to be equal across groups. To test whether metric invariance had been achieved, the resulting χ^2 value was compared to that of the less restrictive “configural” model using the DIFFTEST Mplus command, with a nonsignificant test result suggesting that the additional constraints did not significantly worsen model fit. As the $\Delta\chi^2$ value has been shown to be sensitive to sample size, we also examined the Δ CFI criterion; if the deterioration in CFI for the more constrained model did not exceed 0.01, we considered that level of invariance to be achieved (F. F. Chen, 2007; Cheung & Rensvold, 2002). Finally, to test scalar (or “strong”) invariance, we further constrained item intercepts (for IC items) and thresholds (for LPB items), again comparing this model's fit with the preceding metric model. Where scalar invariance is met, groups can be compared on their mean latent factor scores.

Step 3: Criterion validity. We tested the criterion validity of the latent factors that emerged from the best-fitting model by examining bivariate correlations with external correlates, via standardized ($M = 0$; $SD = 1$) factor scores for early exposure to harsh and warm parenting, and childhood psychiatric comorbidities (externalizing disorder, internalizing disorder, and social-cognitive difficulties). Differences between dependent correlations were tested using Fisher's r -to- z transformation.

Results

Step 1: Model Fitting

Testing the association between IC and LPB total scores in advance of model fitting, we note a moderate correlation ($r = .49$, $p < .001$), implying that there was at least some unique variance underlying the scales. Using the complete pool of 11 items, we estimated (a) unidimensional, (b) two-factor, (c) higher-order, and (d) bifactor structures using CFA techniques. Fit statistics, presented in Table 1, favored the bifactor solution, $\chi^2(33) = 203.35$, $p < .0001$; CFI = .99; TLI = .98; RMSEA = .031; 90% CI [.027, .035] over the unidimensional, two-factor, and higher-order models. It should be noted that model fit information for the higher-order model was identical to that of the two-factor model. Given that only two first-order factors were estimated, the higher-order factor here simply represented an alternative way of capturing covariance between the two factors. Standardized item loadings for this solution are presented in Figure 1 (for factor structures of

Table 1
Model Fit Information for Estimated CFA Models

Model	Model fit statistics				
	χ^2 (df)	p-value	CFI	TLI	RMSEA [90% CIs]
IC and LPB items					
Unidimensional	2,870.45 (44)	<.0001	.83	.79	.108 [.105, .112]
Two-factor (correlated)	665.84 (43)	<.0001	.96	.95	.051 [.048, .055]
Higher-order	665.84 (43)	<.0001	.96	.95	.051 [.048, .055]
Bifactor	203.35 (33)	<.0001	.99	.98	.031 [.027, .035]
Positively and negatively worded items					
Two-factor (correlated)	2,390.04 (43)	<.0001	.86	.82	.100 [.097, .103]
Three-factor (correlated)	449.97 (41)	<.0001	.97	.97	.043 [.039, .046]

Note. IC = interpersonal callousness; LPB = low prosocial behavior; CFI = comparative fit index (good fit $\geq .95$); TLI = Tucker-Lewis index (good fit $\geq .95$); RMSEA = root mean square error of approximation (close fit < .05); CIs = confidence intervals.

alternative models, see Supplemental Figure S2). To test whether this bifactor model improved fit compared with the next best-fitting solution (i.e., two-factor), we tested chi-square differences between these nested models. The significant result, $\Delta\chi^2(10) = 379.20$, $p < .0001$, suggested that the bifactor model improved fit compared with the two-factor model.

The bifactor model featured a general factor (termed IC/LPB), which explained the correlations between all items, and specific IC and LPB factors accounting for residual covariance among their item subsets. All 11 items loaded significantly and positively on IC/LPB (range: .40–.73). With regard to the IC items, although all six loaded significantly on both the general and specific factor, loadings on the IC factor (–.08–.35) were notably weaker than on IC/LPB (.40–.73), suggesting that the general factor accounted well for the variance of these items. In contrast, the five LPB items loaded similarly on the general and specific LPB factor, suggesting that their variances were somewhat split between these factors. All item loadings on the LPB factor were significant and positive (.48–.65).

It was noted that the three IC items with the highest loadings (.30–.35) on the specific IC factor (“good first expression later seen through;” “shallow, fast-changing emotions;” “cold-blooded/callous”) were positively worded, as opposed to the other eight items, which were negatively worded (and subsequently reverse-scored). In order to examine the potential impact of item response biases on factor loadings, we tested two models distinguished by the wording direction of items. First, we estimated a correlated two-factor model that separated positively and negatively worded items (see Supplemental Figure S2, panel C). All fit indices were below accepted minimum thresholds (see Table 1). Second, we tested a three-factor solution that separated positively and negatively worded IC items, along with an LPB factor (see Supplemental Figure S2, panel D). Although an improvement on the two-factor wording model, $\Delta\chi^2(2) = 1,021.35$, $p < .0001$, this three-factor structure still offered a worse fit for the data compared to the bifactor solution, $\Delta\chi^2(8) = 206.44$, $p < .0001$. Another potential source of method effects for our bifactor model was the scaling difference between 5-point IC items and 3-point LPB items. As a sensitivity analysis, we recoded IC to a 3-point scale that roughly corresponded to the LPB items’ response options (*not at all, rarely/sometimes, often/always*). Re-estimating CFA models

with these recoded data, the bifactor solution again yielded the best fit (see Supplemental Table S1). Item loadings for this recoded bifactor model were similar to those of the initial “mixed-scale” solution (see Supplemental Figure S3), although two items (“doesn’t keep promises;” “nongenuine emotional expression”) were now classified as nonsignificant based on two-tailed p values.

Given recent methodological concerns in response to the proliferation of bifactor models (Bonifay et al., 2017; Reise et al., 2016), additional bifactor-specific fit indices are presented in Table 2. First, examining model-based reliability estimates, based on the omega coefficient ($\omega = .88$), a high proportion of the variance in the total score was attributable to all common sources of variance (i.e., general and specific factors). Reliabilities for the IC ($\omega_S = .78$) and LPB subscale scores ($\omega_S = .85$) were likewise high. For the IC/LPB factor, the ω_H statistic indicated that it independently accounted for 70.6% of variance in the total score, while the H value (.84) suggested that this general factor was well-defined by the items. Having partitioned out this variance for the IC/LPB factor, the ω_{HS} score (.07) for the specific IC factor was substantially lower than its original ω score (.78). This suggested that most of the variance explained by this subscale was attributable to the general factor, such that a subscale score based on this latent factor would only explain 6.6% of total score variance. Furthermore, this factor did not reliably represent its underlying items ($H = .27$). Consequently, a latent IC variable would likely possess too little true score variance to enable clinical interpretation (Rodriguez, Reise, & Haviland, 2016b). On the other hand, ω_{HS} (.47) and H values (.69) for the specific LPB factor fell just short of their respective benchmarks for interpretation, as suggested by Reise, Bonifay, and Haviland (2013). Therefore, the latent LPB factor appeared to uniquely explain some proportion of total variance over and above the general factor.

Evaluating the relative unidimensionality of the data, neither the ECV (.64) nor PUC (.55) values for the model reached their accepted thresholds (i.e., $>.70$). This suggested that the data contained some degree of multidimensionality, whereby fitting the data on a single latent factor could introduce unwanted bias to the factor loadings. Moreover, the average relative parameter bias across all items was 19.19%, which exceeded the standard 10%–15% threshold for acceptable bias. However, when calculated

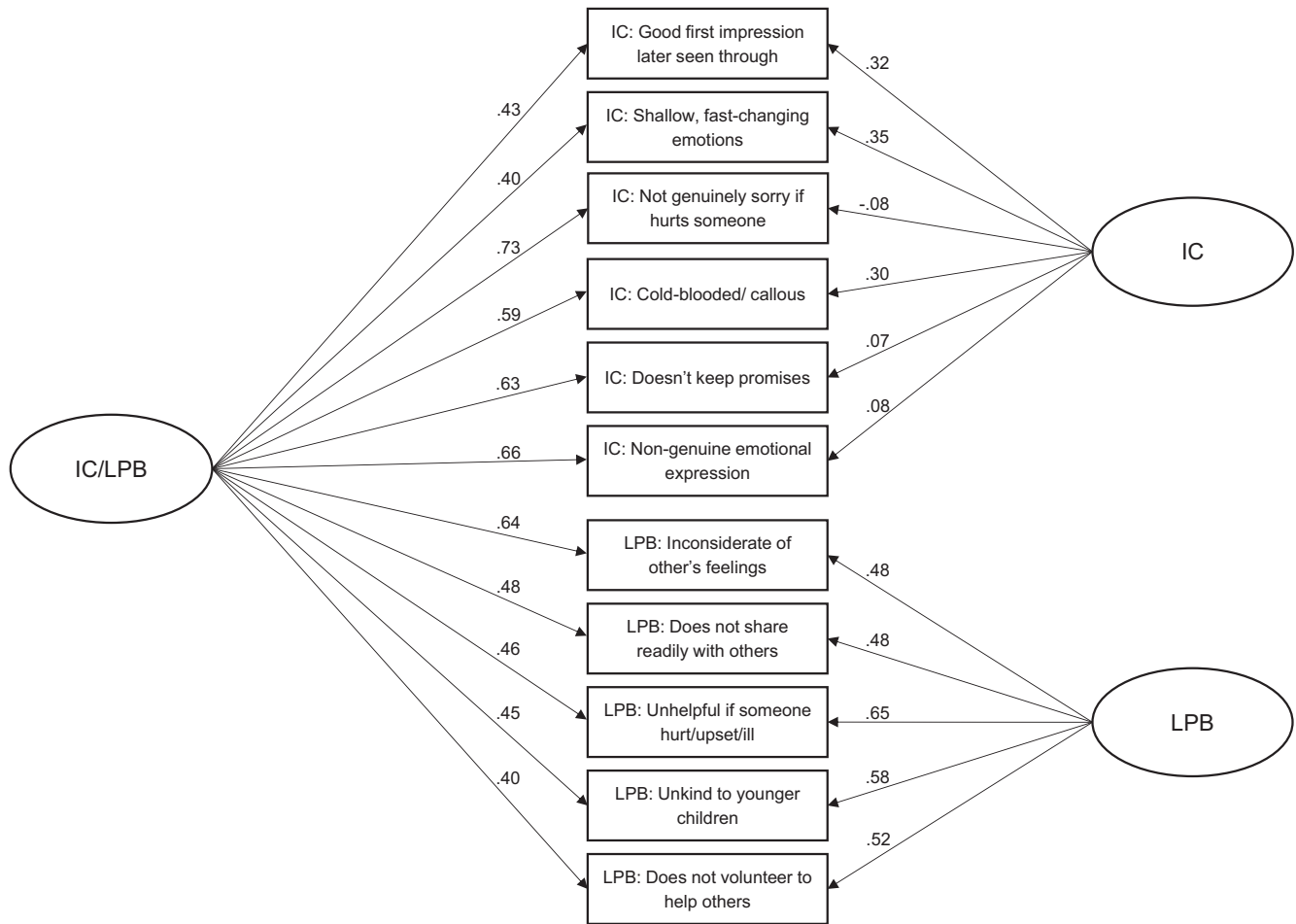


Figure 1. Bifactor model with standardized factor loadings. Solid lines denote significance at $p < .05$. IC = interpersonal callousness; LPB = low prosocial behavior.

separately for IC and LPB, the average parameter bias between unidimensional and bifactor loadings for IC items was within the acceptable range (6.42%), while the average bias for LPB items far exceeded the threshold (34.53%). This, coupled with subthreshold ECV and PUC values for the general factor, offered further support for the multidimensional nature of LPB items compared to IC, as it showed that forcing LPB items into a unidimensional structure

would introduce substantial bias to the factor loadings. This is reflected in the relatively poor fit of a unidimensional model for these data (see Table 1).

Step 2: Sex Differences

Descriptive statistics for manifest IC and LPB total scores, across the entire sample and separately for males and females, are presented in Table 3. Males and females did not differ on mean IC scores, $t(5,446) = -0.86$, $p = .39$, $d = .02$, in line with findings from previous studies in ALSPAC that utilize this scale (e.g., Barker et al., 2011; Barker & Salekin, 2012). However, males showed significantly higher mean LPB scores compared with females, although the effect size here was small, $t(5,446) = 10.68$, $p < .001$, $d = .29$.

With regard to the bifactor model itself, the model provided a good fit when estimated separately for males, $\chi^2(33) = 122.98$, $p < .0001$; CFI = .99; TLI = .98; RMSEA = .032; 90% CI [.026, .038], and females, $\chi^2(33) = 121.63$, $p < .0001$; CFI = .99; TLI = .98; RMSEA = .031; 90% CI [.026, .038]. Results of successive tests of increasing multiple-group measurement invariance are presented in Table 4. All levels of invariance, up to scalar invariance, could be assumed, as evidenced by nonsignificant decreases

Table 2
Additional Fit Indices for the Estimated Bifactor Model

Factor	ω/ω_S	Bifactor-derived statistic			
		ω_H/ω_{HS}	H	ECV	PUC
IC/LPB	.878	.706	.839	.644	.545
IC	.780	.066	.269		
LPB	.851	.470	.686		

Note. ω = omega (i.e., for IC/LPB); ω_S = omega subscale (i.e., for IC and LPB); ω_H = omega hierarchical (for IC/LPB); ω_{HS} = omega hierarchical subscale (for IC and LPB); H = construct replicability; ECV = explained common variance; PUC = percentage uncontaminated correlations.

Table 3
Descriptive Statistics and Sex Differences for IC and LPB Scores

Scale	M (SD)			Sex differences		
	Total sample (N = 5,463)	Males (n = 2,735)	Females (n = 2,713)	t(5,446)	p-value	Cohen's d
IC (range: 0–24)	4.68 (3.13)	4.64 (3.12)	4.71 (3.15)	-.86	.390	.02
LPB (range: 0–10)	1.72 (1.70)	1.97 (1.78)	1.48 (1.59)	10.68	<.001	.29

Note. IC = interpersonal callousness; LPB = low prosocial behavior. Discrepancy between total and sex-specific samples is due to 12 participants with missing information for sex.

in χ^2 and CFI values between each more constrained model and the less restrictive model preceding it.

Step 3: Criterion Validity

Correlations between latent factors from the bifactor model and the five criterion variables are presented in Table 5. External correlates were captured using regression-based factor scores derived from separate CFAs (for the underlying latent factor structures of these scores, see Supplemental Figure S1). Given the apparent unreliability of the latent IC factor indicated by the bifactor-specific fit statistics, we did not interpret associations involving this factor; these are available in Supplemental Table S2. Item loadings for the IC/LPB and LPB factors, meanwhile, remained broadly similar to the initial CFA model once criterion variables had been entered into a multivariate model in order to examine these associations (see Supplemental Figure S4).

The general IC/LPB factor was significantly associated with higher levels of harsh parenting, $r = .20, p < .001$; externalizing disorder, $r = .54, p < .001$; internalizing disorder, $r = .19, p < .001$; social-cognitive difficulties, $r = .54, p < .001$; and lower levels of warm parenting, $r = -.16, p < .001$. The specific LPB factor was significantly associated with higher levels of externalizing disorder, $r = .04, p = .03$; social-cognitive difficulties, $r = .19, p < .001$; and lower levels of warm parenting, $r = -.07, p = .001$. Testing differences between significant correlations for the two factors, the correlations for IC/LPB with harsh parenting ($z = 5.18, p < .001$), externalizing disorder ($z = 29.23, p < .001$), and social-cognitive difficulties ($z = 26.61, p < .001$) were significantly greater than those for LPB.

Discussion

Using a longitudinal birth cohort, the present study compared several alternative models that captured the shared and unique variances underlying IC and LPB. A bifactor model offered the

best fit for these data, with the resulting general and specific factors presenting distinct associations with external measures of early parenting and psychiatric comorbidity. We highlight three main findings here.

First, similar to previous childhood studies of youth callousness (Dadds et al., 2005; Viding et al., 2005), we identified substantial shared variance between IC and LPB constructs. More specifically, our best-fitting model was a bifactor solution with two reliable orthogonal factors: a general (IC/LPB) factor and a residual LPB factor. Invariance testing also suggested that the latent factor structure for this bifactor model was consistent between males and females. All 11 items loaded significantly on IC/LPB, and most loaded higher here than the residual factors. In addition, the general factor accounted for the majority (70.6%) of total score variance. Based on bivariate associations, IC/LPB represented a more severe psychiatric phenotype than the unique residual variances for IC or LPB, showing the widest range of associations with risk factors and comorbidities. Specifically, regarding early parenting, IC/LPB was uniquely associated with greater harshness and less warmth, while for comorbidities, IC/LPB was associated with higher levels of externalizing and internalizing problems and social-cognitive difficulties. Overall, this general factor may align with the combined "callous-low prosocial" measures previously derived in childhood using APSD and SDQ items (Dadds et al., 2005; Viding et al., 2005). Like these studies, we also measured LPB using the SDQ, while the IC items in ALSPAC have themselves correlated highly ($r = .81$) with the APSD's Callous-Unemotional subscale (Moran et al., 2009).

Second, the viability of a specific IC factor was not supported by further interrogation of the model using bifactor-specific fit statistics. These indicated that the factor was an unrepresentative and unreliable estimate of the variance underlying the IC items, and it was therefore unlikely that the associations between this factor and external correlates would be trustworthy or replicable (Rodriguez et al., 2016b). Indeed, many of the observed associations were

Table 4
Multiple-Group Measurement Invariance in the Bifactor Model Between Males and Females

Model	Model fit statistics					Difference testing			
	χ^2 (df)	p-value	CFI	TLI	RMSEA [90% CIs]	$\Delta\chi^2$	Δdf	p-value	ΔCFI
Configural	318.42 (74)	<.0001	.985	.978	.035 [.031, .039]				
Metric/weak	220.81 (91)	<.0001	.992	.990	.023 [.019, .027]	27.23	17	.055	.007
Scalar/strong	197.31 (96)	<.0001	.994	.993	.020 [.016, .024]	3.16	5	.677	.002

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CIs = confidence intervals.

Table 5
Correlations Among Estimated IC/LPB and LPB Factors, Early Parenting Styles, and Psychiatric Comorbidities

Variable	1	2	3	4	5	6	7
1. IC/LPB	1						
2. LPB	.000	1					
3. Harsh parenting (age 2–4)	.203*** ^a	.000 ^b	1				
4. Warm parenting (age 2–4)	-.164*** ^a	-.066*** ^b	-.144***	1			
5. Externalizing disorder (age 7–13)	.535*** ^a	.035*** ^b	.245***	-.145***	1		
6. Internalizing disorder (age 7–13)	.194*** ^a	.001 ^b	.066***	-.039*	.314***	1	
7. Social-cognitive difficulties (age 7–13)	.538*** ^a	.086*** ^b	.192***	-.129***	.753***	.340***	1

Note. $N = 5,463$. For IC/LPB and LPB (i.e., columns 1 and 2), correlations with different superscripts are significantly different from one another across rows, based on tests of equality of dependent correlations. Correlations for the latent IC factor can be found in Supplemental Table S2.

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

rather counterintuitive (see Supplemental Table S2). It appears that once shared variance with LPB is partitioned out (via IC/LPB), the remaining IC variance does not offer a unique contribution over and above this general factor, such that IC in the absence of LPB may not represent a substantively meaningful construct, at least within this young sample. Practically, it may be that very few children or adolescents with high levels of IC would not also, to an extent, also show low levels of prosocial expression. Alternatively, given that the highest-loading items on the IC factor were the only positively worded items in the entire pool, it could be that this factor simply captures measurement variance based on individual differences in responding to differently worded items, as suggested by item response analyses of the factor structures identified in other psychopathy scales (Ray et al., 2016).

Third, LPB appeared to represent a more multidimensional construct compared with IC. Items loaded similarly on specific LPB and IC/LPB factors, and bifactor-specific reliability statistics for the specific factor were approaching their suggested thresholds. Additionally, ECV, PUC, and parameter bias values did not fully support a unidimensional structure in which the IC/LPB factor would sufficiently explain the variance in the data. However, once variance shared with IC was accounted for, the remaining variance for LPB did not seem to reflect as severe a profile of risk and impairment, as it was only weakly related to greater externalizing problems and social-cognitive difficulty, and lower maternal warmth. Our findings may suggest that this residual LPB variance represents a phenotype characterized by social-cognitive impairments associated with autism, given that this measure of social cognition is frequently used to inform ASD diagnosis (Skuse et al., 2009). Lower prosocial behavior, as measured using the SDQ, has previously been reported in youth with ASD (Russell et al., 2012). ASD-related traits have previously been linked with poor cognitive—but not affective—empathy, in contrast to callousness, where the opposite profile is observed (e.g., Pasalich et al., 2014; Schwenck et al., 2012). Thus, in tapping deficits in social communication, the variance captured by the residual LPB factor may reflect an inability to understand the inner states of others (cognitive empathy), rather than a lack of concern for them (affective empathy). This distinction may account for the largely unique profiles of etiological influences between autistic and callous traits within twin samples (O’Nions et al., 2015). Future work should aim to further unpack the phenotypic profile of this residual factor,

by examining associations with cognitive and affective components of empathy.

Clinical Implications

It has been suggested that current clinical conceptualizations of callousness, such as the “limited prosocial emotions” specifier for CD in DSM-5 (American Psychiatric Association, 2013), treat the *presence* of callousness and *absence* of prosocial expression as conceptually equivalent (Willoughby et al., 2015). In support of this, we identified considerable overlap between IC and LPB, to the extent that their shared variance (IC/LPB) accounted for most of the meaningful variance for IC. Therefore, LPB items may be useful to supplement and expand existing characteristics of youth IC traits, and better inform diagnostic efforts. These findings suggest that the broader construct of psychopathy-like symptoms, as well as its underlying dimensions, could be helpful in understanding youth with CP and allowing for more tailored treatment efforts (Salekin, 2016a, 2016b). In general, the presence of IC has been associated with poor outcomes in parent-training interventions aimed at reducing CP (for review, see D. J. Hawes et al., 2014). However, a recent school-based intervention aiming to promote prosocial behavior in adolescence has led to decreases in aggression, and increases in academic achievement, at 6-month follow-up (Caprara et al., 2014). Therefore, given the apparent phenotypic overlap with IC traits, targeting (low) prosocial expression and behavior may offer a novel treatment target to help mitigate the more severe patterns of maladjustment generally conferred by IC.

Strengths and Limitations

The current study benefitted from a large sample size, prospective design, and use of a latent variable framework to combine repeated measures of external correlates, in an effort to minimize potential measurement error. Nonetheless, several limitations must be acknowledged. First, IC and LPB were only available once at the same time-point (age 13). This precluded examination of the stability of the bifactor model, or whether the best-fitting factor structure for IC and LPB items changed across development (i.e., the degree of age invariance for this model). Second, our community-based sample may present a more limited range of

severity for psychiatric and behavioral dysfunction, and the extent to which these results apply to more high-risk or clinical populations is unclear. For example, estimating a model among high-CP youth may yield a different factor structure or profile of association with included risks and comorbidities. Third, our IC and LPB measures, and our external correlates, were all drawn from parent reports, introducing the possibility of shared method variance. Future attempts to replicate these results should draw on multiple informants where possible (e.g., self- and teacher-report). Fourth, although we have, to the best of our efforts, attempted to test comparative models and/or conduct sensitivity analyses around differences in response scales and wording direction among our 11 items, we cannot conclusively rule out the possible influence of method effects on our residual factors under the current analytic framework. Fifth, and finally, although IC traits have been consistently associated with high heritability in twin designs (e.g., Viding et al., 2005), the extent to which the general IC/LPB factor reflects a common genetic liability underlying IC and LPB, or simply high phenotypic association between these items, could not be examined using the current data. Genetically informative designs, such as twin studies, are needed to compare relative genetic and environmental influences for IC traits and (low) prosocial behavior, and determine whether IC/LPB captures common variance at the genotypic, as well as phenotypic, level.

Conclusions

Extending previous studies that report an association between IC and LPB (Barker et al., 2011; Meehan et al., 2017), the current study identifies substantial shared variance underlying these two constructs in early adolescence and demonstrates that this general factor designates a more severe profile of childhood environmental risk and co-occurring psychopathology. In contrast, it appears that the residual variance for LPB is mainly characterized by social-cognitive impairment. These findings indicate that high levels of IC traits predominantly also involve low levels of prosocial behavior, such that the inclusion of LPB items may enhance measurement of IC. At a broader level, findings suggest that capturing facets of psychopathy beyond affective (i.e., callous-unemotional) traits alone, such as more interpersonal features, might enable better diagnostic systems and more effective treatment of childhood conduct problems. More generally, the fact that bifactor-specific reliability estimates showed that a specific IC factor was not informative over and above the general factor highlights the importance of generating these additional indices for superior-fitting bifactor models, in order to determine whether data are genuinely multidimensional, and whether all factors reflect reliable and meaningful variance.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Barker, E. D., Oliver, B. R., Viding, E., Salekin, R. T., & Maughan, B. (2011). The impact of prenatal maternal risk, fearless temperament and early parenting on adolescent callous-unemotional traits: A 14-year longitudinal investigation. *Journal of Child Psychology and Psychiatry*, 52, 878–888. <http://dx.doi.org/10.1111/j.1469-7610.2011.02397.x>
- Barker, E. D., & Salekin, R. T. (2012). Irritable oppositional defiance and callous unemotional traits: Is the association partially explained by peer victimization? *Journal of Child Psychology and Psychiatry*, 53, 1167–1175. <http://dx.doi.org/10.1111/j.1469-7610.2012.02579.x>
- Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5, 184–186. <http://dx.doi.org/10.1177/2167702616657069>
- Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., . . . Davey Smith, G. (2013). Cohort profile: The 'Children of the 90s'—The index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42, 111–127. <http://dx.doi.org/10.1093/ije/dys064>
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Byrd, A. L., Hawes, S. W., Loeber, R., & Pardini, D. A. (2018). Interpersonal callousness from childhood to adolescence: Developmental trajectories and early risk factors. *Journal of Clinical Child and Adolescent Psychology*, 47, 467–482. <http://dx.doi.org/10.1080/15374416.2016.1144190>
- Byrd, A. L., Loeber, R., & Pardini, D. A. (2012). Understanding desisting and persisting forms of delinquency: The unique contributions of disruptive behavior disorders and interpersonal callousness. *Journal of Child Psychology and Psychiatry*, 53, 371–380. <http://dx.doi.org/10.1111/j.1469-7610.2011.02504.x>
- Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, 11, 302–306. <http://dx.doi.org/10.1111/1467-9280.00260>
- Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zuffianò, A., Alessandri, G., Vecchio, G., . . . Bridglall, B. (2014). Positive effects of promoting prosocial behavior in early adolescence. *International Journal of Behavioral Development*, 38, 386–396. <http://dx.doi.org/10.1177/0165025414531464>
- Carlo, G., Mestre, M. V., McGinley, M. M., Tur-Porcar, A., Samper, P., & Opal, D. (2014). The protective role of prosocial behaviors on antisocial behaviors: The mediating effects of deviant peer affiliation. *Journal of Adolescence*, 37, 359–366. <http://dx.doi.org/10.1016/j.adolescence.2014.02.009>
- Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2010). Feelings or cognitions? Moral cognitions and emotions as longitudinal predictors of prosocial and aggressive behaviors. *Personality and Individual Differences*, 48, 872–877. <http://dx.doi.org/10.1016/j.paid.2010.02.010>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. <http://dx.doi.org/10.1080/10705510701301834>
- Chen, X., Liu, M., Rubin, K. H., Cen, G., Gao, X., & Li, D. (2002). Sociability and prosocial orientation as predictors of youth adjustment: A seven-year longitudinal study in a Chinese sample. *International Journal of Behavioral Development*, 26, 128–136. <http://dx.doi.org/10.1080/01650250042000690>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- 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. <http://dx.doi.org/10.1037/0022-006X.73.3.400>
- Dadds, M. R., Perry, Y., Hawes, D. J., Merz, S., Riddell, A. C., Haines, D. J., . . . Abeygunawardane, A. I. (2006). Attention to the eyes and

- fear-recognition deficits in child psychopathy. *The British Journal of Psychiatry*, 189, 280–281. <http://dx.doi.org/10.1192/bjp.bp.105.018150>
- Day, R. D., & Padilla-Walker, L. M. (2009). Mother and father connectedness and involvement during early adolescence. *Journal of Family Psychology*, 23, 900–904. <http://dx.doi.org/10.1037/a0016438>
- Edens, J. F., Skopp, N. A., & Cahill, M. A. (2008). Psychopathic features moderate the relationship between harsh and inconsistent parental discipline and adolescent antisocial behavior. *Journal of Clinical Child and Adolescent Psychology*, 37, 472–476. <http://dx.doi.org/10.1080/15374410801955938>
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg & W. Damon (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 646–718). New York, NY: Wiley.
- Eisenberg, N., Spinrad, T. L., & Knafo-Noam, A. (2015). Prosocial development. In J. E. Lamb & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (7th ed., pp. 610–656). New York, NY: Wiley.
- Eivers, A. R., Brendgen, M., Vitaro, F., & Borge, A. I. (2012). Concurrent and longitudinal links between children's and their friends' antisocial and prosocial behavior in preschool. *Early Childhood Research Quarterly*, 27, 137–146. <http://dx.doi.org/10.1016/j.ecresq.2011.05.001>
- Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment*, 13, 454–469. <http://dx.doi.org/10.1177/1073191106287354>
- Flynn, E., Ehrenreich, S. E., Beron, K. J., & Underwood, M. K. (2015). Prosocial behavior: Long-term trajectories and psychosocial outcomes. *Social Development*, 24, 462–482. <http://dx.doi.org/10.1111/sode.12100>
- Fontaine, N. M., McCrory, E. J., Boivin, M., Moffitt, T. E., & Viding, E. (2011). Predictors and outcomes of joint trajectories of callous-unemotional traits and conduct problems in childhood. *Journal of Abnormal Psychology*, 120, 730–742. <http://dx.doi.org/10.1037/a0022620>
- Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., . . . Lawlor, D. A. (2013). Cohort profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, 42, 97–110. <http://dx.doi.org/10.1093/ije/dys066>
- Frick, P. J. (2004). *Inventory of callous-unemotional traits*. Unpublished rating scale. University of New Orleans, New Orleans, LA.
- Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment*, 12, 382–393. <http://dx.doi.org/10.1037/1040-3590.12.4.382>
- Frick, P. J., & Hare, R. D. (2001). *The antisocial process screening device*. Toronto, Ontario, Canada: Multi-Health Systems.
- Frick, P. J., O'Brien, B. S., Wootton, J. M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology*, 103, 700–707. <http://dx.doi.org/10.1037/0021-843X.103.4.700>
- 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. <http://dx.doi.org/10.1037/a0033076>
- Gerbino, M., Zuffiano, A., Eisenberg, N., Castellani, V., Luengo Kanacri, B. P., Pastorelli, C., & Caprara, G. V. (2018). Adolescents' prosocial behavior predicts good grades beyond intelligence and personality traits. *Journal of Personality*, 86, 247–260. <http://dx.doi.org/10.1111/jopy.12309>
- Goodman, A., Heiervang, E., Collishaw, S., & Goodman, R. (2011). The 'DAWBA bands' as an ordered-categorical measure of child mental health: Description and validation in British and Norwegian samples. *Social Psychiatry and Psychiatric Epidemiology*, 46, 521–532. <http://dx.doi.org/10.1007/s00127-010-0219-x>
- Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, 38, 1179–1191. <http://dx.doi.org/10.1007/s10802-010-9434-x>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586. <http://dx.doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41, 645–655. <http://dx.doi.org/10.1111/j.1469-7610.2000.tb02345.x>
- Hare, R. D. (2003). *The Hare Psychopathy Checklist—revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.
- Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment*, 1, 6–17. <http://dx.doi.org/10.1037/1040-3590.1.1.6>
- Hastings, P. D., Utendale, W. T., & Sullivan, C. (2007). The socialization of prosocial development. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization* (pp. 638–664). New York, NY: Guilford Press.
- Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical Child and Family Psychology Review*, 17, 248–267. <http://dx.doi.org/10.1007/s10567-014-0167-1>
- Hawes, S. W., Byrd, A. L., Henderson, C. E., Gazda, R. L., Burke, J. D., Loeber, R., & Pardini, D. A. (2014). Refining the parent-reported Inventory of Callous-Unemotional Traits in boys with conduct problems. *Psychological Assessment*, 26, 256–266. <http://dx.doi.org/10.1037/a0034718>
- Hipwell, A. E., Pardini, D. A., Loeber, R., Semboer, M., Keenan, K., & Stouthamer-Loeber, M. (2007). Callous-unemotional behaviors in young girls: Shared and unique effects relative to conduct problems. *Journal of Clinical Child and Adolescent Psychology*, 36, 293–304. <http://dx.doi.org/10.1080/15374410701444165>
- Holinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54. <http://dx.doi.org/10.1007/BF02287965>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Jameel, L., Vyas, K., Bellesi, G., Roberts, V., & Channon, S. (2014). Going 'above and beyond': Are those high in autistic traits less pro-social? *Journal of Autism and Developmental Disorders*, 44, 1846–1858. <http://dx.doi.org/10.1007/s10803-014-2056-3>
- Knafo, A., & Plomin, R. (2006). Parental discipline and affection and children's prosocial behavior: Genetic and environmental links. *Journal of Personality and Social Psychology*, 90, 147–164. <http://dx.doi.org/10.1037/0022-3514.90.1.147>
- Kokko, K., & Pulkkinen, L. (2000). Aggression in childhood and long-term unemployment in adulthood: A cycle of maladaptation and some protective factors. *Developmental Psychology*, 36, 463–472. <http://dx.doi.org/10.1037/0012-1649.36.4.463>
- Kokko, K., Tremblay, R. E., Lacourse, E., Nagin, D. S., & Vitaro, F. (2006). Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. *Journal of Research on Adolescence*, 16, 403–428. <http://dx.doi.org/10.1111/j.1532-7795.2006.00500.x>

- Kroneman, L. M., Hipwell, A. E., Loeber, R., Koot, H. M., & Pardini, D. A. (2011). Contextual risk factors as predictors of disruptive behavior disorder trajectories in girls: The moderating effect of callous-unemotional features. *Journal of Child Psychology and Psychiatry*, 52, 167–175. <http://dx.doi.org/10.1111/j.1469-7610.2010.02300.x>
- Meehan, A. J., Maughan, B., Cecil, C. A. M., & Barker, E. D. (2017). Interpersonal callousness and co-occurring anxiety: Developmental validity of an adolescent taxonomy. *Journal of Abnormal Psychology*, 126, 225–236. <http://dx.doi.org/10.1037/abn0000235>
- Moran, P., Ford, T., Butler, G., & Goodman, R. (2008). Callous and unemotional traits in children and adolescents living in Great Britain. *The British Journal of Psychiatry*, 192, 65–66. <http://dx.doi.org/10.1192/bjp.bp.106.034876>
- 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. <http://dx.doi.org/10.1097/CHI.0b013e3181b766ab>
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide: 1998–2012* (7th ed.). Los Angeles, CA: Author.
- Nantel-Vivier, A., Pihl, R. O., Côté, S., & Tremblay, R. E. (2014). Developmental association of prosocial behaviour with aggression, anxiety and depression from infancy to preadolescence. *Journal of Child Psychology and Psychiatry*, 55, 1135–1144. <http://dx.doi.org/10.1111/jcpp.12235>
- Office of Population Censuses and Surveys. (1991). *Standard occupational classification* (Vol. 3). London, UK: HMSO.
- O'Nions, E., Tick, B., Rijdsdijk, F., Happe, F., Plomin, R., Ronald, A., & Viding, E. (2015). Examining the genetic and environmental associations between autistic social and communication deficits and psychopathic callous-unemotional traits. *PLoS ONE*, 10, e0134331. <http://dx.doi.org/10.1371/journal.pone.0134331>
- Oxford, M., Cavell, T. A., & Hughes, J. N. (2003). Callous/unemotional traits moderate the relation between ineffective parenting and child externalizing problems: A partial replication and extension. *Journal of Clinical Child and Adolescent Psychology*, 32, 577–585. http://dx.doi.org/10.1207/S15374424JCCP3204_10
- Pardini, D. A., Lochman, J. E., & Powell, N. (2007). The development of callous-unemotional traits and antisocial behavior in children: Are there shared and/or unique predictors? *Journal of Clinical Child and Adolescent Psychology*, 36, 319–333. <http://dx.doi.org/10.1080/15374410701444215>
- Pardini, D. A., 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 and Adolescent Psychology*, 35, 46–59. http://dx.doi.org/10.1207/s15374424jccp3501_5
- Pardini, D. A., Stepp, S., Hipwell, A., Stouthamer-Loeber, M., & Loeber, R. (2012). The clinical utility of the proposed *DSM-5* callous-unemotional subtype of conduct disorder in young girls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51, 62–73. <http://dx.doi.org/10.1016/j.jaac.2011.10.005>
- Pasalich, D. S., Dadds, M. R., & Hawes, D. J. (2014). Cognitive and affective empathy in children with conduct problems: Additive and interactive effects of callous-unemotional traits and autism spectrum disorders symptoms. *Psychiatry Research*, 219, 625–630. <http://dx.doi.org/10.1016/j.psychres.2014.06.025>
- Pasalich, D. S., Dadds, M. R., Hawes, D. J., & Brennan, J. (2011). Do callous-unemotional traits moderate the relative importance of parental coercion versus warmth in child conduct problems? An observational study. *Journal of Child Psychology and Psychiatry*, 52, 1308–1315. <http://dx.doi.org/10.1111/j.1469-7610.2011.02435.x>
- Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2016). Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological Assessment*, 28, 394–404. <http://dx.doi.org/10.1037/pas0000183>
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95, 129–140. <http://dx.doi.org/10.1080/00223891.2012.725437>
- Reise, S. P., Kim, D. S., Mansolf, M., & Widaman, K. F. (2016). Is the bifactor model a better model or is it just better at modeling implausible responses? Application of iteratively reweighted least squares to the Rosenberg Self-Esteem Scale. *Multivariate Behavioral Research*, 51, 818–838. <http://dx.doi.org/10.1080/00273171.2016.1243461>
- Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17, 354–373. <http://dx.doi.org/10.1037/a0029315>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98, 223–237. <http://dx.doi.org/10.1080/00223891.2015.1089249>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21, 137–150. <http://dx.doi.org/10.1037/met0000045>
- Rowe, R., Maughan, B., Moran, P., Ford, T., Briskman, J., & Goodman, R. (2010). The role of callous and unemotional traits in the diagnosis of conduct disorder. *Journal of Child Psychology and Psychiatry*, 51, 688–695. <http://dx.doi.org/10.1111/j.1469-7610.2009.02199.x>
- Russell, G., Golding, J., Norwich, B., Emond, A., Ford, T., & Steer, C. (2012). Social and behavioural outcomes in children diagnosed with autism spectrum disorders: A longitudinal cohort study. *Journal of Child Psychology and Psychiatry*, 53, 735–744. <http://dx.doi.org/10.1111/j.1469-7610.2011.02490.x>
- Salekin, R. T. (2016a). Psychopathy in childhood: Toward better informing the DSM-5 and ICD-11 conduct disorder specifiers. *Personality Disorders*, 7, 180–191. <http://dx.doi.org/10.1037/per0000150>
- Salekin, R. T. (2016b). Psychopathy in childhood: Why should we care about grandiose-manipulative and daring-impulsive traits? *British Journal of Psychiatry*, 209, 189–191. <http://dx.doi.org/10.1192/bjp.bp.115.179051>
- Salekin, R. T. (2017). Research review: What do we know about psychopathic traits in children? *Journal of Child Psychology and Psychiatry*, 58, 1180–1200. <http://dx.doi.org/10.1111/jcpp.12738>
- Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., . . . Freitag, C. M. (2012). Empathy in children with autism and conduct disorder: Group-specific profiles and developmental aspects. *Journal of Child Psychology and Psychiatry*, 53, 651–659. <http://dx.doi.org/10.1111/j.1469-7610.2011.02499.x>
- Skuse, D. H., Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits: Heritability, reliability and validity of the Social and Communication Disorders Checklist. *The British Journal of Psychiatry*, 187, 568–572. <http://dx.doi.org/10.1192/bjp.187.6.568>
- Skuse, D. H., Mandy, W., Steer, C., Miller, L. L., Goodman, R., Lawrence, K., . . . Golding, J. (2009). Social communication competence and functional adaptation in a general population of children: Preliminary evidence for sex-by-verbal IQ differential risk. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48, 128–137. <http://dx.doi.org/10.1097/CHI.0b013e31819176b8>
- Viding, E., Blair, R. J., Moffitt, T. E., & Plomin, R. (2005). Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry*, 46, 592–597. <http://dx.doi.org/10.1111/j.1469-7610.2004.00393.x>
- Viding, E., Simmonds, E., Petrides, K. V., & Frederickson, N. (2009). The contribution of callous-unemotional traits and conduct problems to bul-

- lying in early adolescence. *Journal of Child Psychology and Psychiatry*, 50, 471–481. <http://dx.doi.org/10.1111/j.1469-7610.2008.02012.x>
- Vitaro, F., Brendgen, M., Larose, S., & Tremblay, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology*, 97, 617–629. <http://dx.doi.org/10.1037/0022-0663.97.4.617>
- Waller, R., Gardner, F., Hyde, L. W., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2012). Do harsh and positive parenting predict parent reports of deceitful-callous behavior in early childhood? *Journal of Child Psychology and Psychiatry*, 53, 946–953. <http://dx.doi.org/10.1111/j.1469-7610.2012.02550.x>
- Waller, R., Wright, A. G., Shaw, D. S., Gardner, F., Dishion, T. J., Wilson, M. N., & Hyde, L. W. (2015). Factor structure and construct validity of the parent-reported Inventory of Callous-Unemotional traits among high-risk 9-year-olds. *Assessment*, 22, 561–580. <http://dx.doi.org/10.1177/1073191114556101>
- Watkins, M. W. (2013). Omega [Computer software]. Phoenix, AZ: Ed & Psych Associates.
- Willoughby, M. T., Mills-Koonce, W. R., Waschbusch, D. A., & Gottfredson, N. C. (2015). An examination of the parent report version of the Inventory of Callous-Unemotional Traits in a community sample of first-grade children. *Assessment*, 22, 76–85. <http://dx.doi.org/10.1177/1073191114534886>
- Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British Journal of Psychiatry*, 195, 249–256. <http://dx.doi.org/10.1192/bjp.bp.108.053751>

Received April 24, 2018

Revision received October 1, 2018

Accepted October 16, 2018 ■