Archives of Forensic Psychology
2014, Vol. 1, No. 1, 27–48

ISSN 2334-2749

Establishing Cut-off Scores for the Parent-Reported Inventory of Callous-Unemotional Traits

Eva R. Kimonis*

School of Psychology, Mathews Building, The University of New South Wales, Sydney, NSW 2052, Australia,

e.kimonis@unsw.edu.au

Kostas A. Fanti

Department of Psychology, University of Cyprus, Nicosia, Cyprus

Jay P. Singh

Global Institute of Forensic Research, Reston, VA, USA

Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Faculty of Health Sciences, Molde University College, Molde, Norway

Callous-unemotional (CU) traits (i.e., lack of empathy/guilt/concern for others) have proven useful for identifying a unique subgroup of antisocial youths at risk for severe, persistent, and impairing conduct problems attributed to distinct etiological processes. Several tools for measuring CU traits alone or as part of a broader assessment of psychopathy exist but none have established cut-off scores for making categorical decisions about youth. The aim of the present study was to establish clinically meaningful cut-off scores on the parent-reported Inventory of Callous-Unemotional Traits (ICU) for the purpose of identifying children with high stable co-occurring conduct problems (CP) and CU traits (CP+CU), while balancing costs of false positives and false negatives. Participants included 1,370 school-aged (M_{age} = 9.38, SD_{age} = 1.64 at baseline) boys and girls followed prospectively over 18 months. Several statistical indices were applied to establish optimal cut-off scores for identifying those 2.3% of children on a trajectory of high stable co-occurring CP+CU according to latent class growth analyses. Results indicated that children who scored at or above the identified ICU cut-off scores (24 for mother-report and 27 for father-report) were significantly more likely to engage in future self-reported bullying compared to children who scored below the thresholds. With encouraging evidence for the success of nuanced treatments for children with CP+CU, these findings may assist in screening children that might benefit from them.

Keywords: Callous-unemotional traits, Conduct problems, Assessment, Cut-off scores, Longitudinal

Youth exhibiting significant and impairing patterns of antisocial and aggressive behavior are heterogeneous with respect to their severity, prognosis, and presumed etiology (Frick & Viding, 2009). Further, a substantial body of research suggests that non-normative levels of callous-unemotional (CU) traits (i.e., lack of empathy/guilt/concern for others) are useful for identifying antisocial youth who show a distinct pattern of severe, chronic, and aggressive conduct problems that are resistant to traditional mental health interventions, and thought to be underpinned by distinct causal factors (Frick, Ray, Thornton, & Kahn, 2014). CU traits comprise one dimension of psychopathy along with narcissism/deceitfulness and impulsivity/irresponsibility.

A key challenge in identifying this at-risk population comes in translating continuous scores obtained from existing tools into dichotomous decisions that can inform prognosis and guide treatment planning. Cut-off scores for commonly used measures of CU traits have not yet been established. Although we recognize that the most compelling data suggest that psychopathy is a dimensional trait rather than a categorical taxon (Edens, Marcus, Lilienfeld, & Poythress, 2006; Lilienfeld, 1994; Murrie, Marcus, et al., 2007; cf. Harris, Rice, & Quinsey, 1994; Vasey, Kotov, Frick, & Loney, 2005), using arbitrary cut-off scores (e.g., deviation from the sample mean) to make important decisions may be either over- or under-inclusive. The present study aimed to address this risk and increase fairness in decision-making by empirically identifying objective cut-off scores on one of the most comprehensive measures of CU traits currently available: the Inventory of Callous-Unemotional Traits (ICU). A primary motivation in achieving this aim was to optimize the accurate identification of children with poor prognosis—i.e., on a trajectory of stable high co-occurring conduct problems (CP) and CU traits (CP+CU)—thus carefully balancing the competing risks of identifying too many or too few youths with non-normative CU traits.

Attempts to categorize children on the basis of CU traits, which are thought to signal risk for psychopathy in adulthood, naturally raise important and valid concerns. This criticism is rooted in part in concerns over stigma associated with labeling youth, in part in the limited longitudinal research literature establishing the stability of CU traits from childhood to adulthood, and in part in the potential for harm that comes with the misuse of psychological instruments (Seagrave & Grisso, 2002). With respect to stability, a long-term follow-up study measuring psychopathic (including CU) traits at age 13 years found that the construct was moderately stable to age 24 years ($r = .32$), despite using different informants and assessment instruments across the two age periods (Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007). Of those 13-year olds scoring in the top 20% on the psychopathy measure, 14% went on to warrant a diagnosis of psychopathy in young adulthood, making this construct more stable than other forms of childhood psychopathology (Mash & Dozois, 2003). Furthermore, the 11-year correlation is equivalent to that typically seen when different informants use the same instrument at the same time-point to rate an individual's behavior on a construct. Within any given developmental stage (childhood or adolescence), the general stability of CU or psychopathic traits tends to be higher at around 30% (e.g., Lee, Klaver, Hart, Moretti, & Douglas, 2009).

CU traits also signal risk for other antisocial outcomes in adulthood. For example, after accounting for the severity of co-occurring conduct problems, CU traits present at elementary and middle school age predicted criminal behavior and other antisocial outcomes (e.g., antisocial personality symptoms) in adulthood (Byrd, Loeber, & Pardini,

2012; McMahon, Witkiewitz, & Kotler, 2010). These findings highlight the importance of early identification and treatment of children with CP+CU given their heightened risk for long-term impairment.

With respect to stigma, although there is no research directly testing the effects of labeling youths as having CU traits, there has been research on the negative effects of the use of the term *psychopathy* when applied to children and adolescents (for a review, see Murrie, Boccaccini, McCoy, & Cornell, 2007). The term *psychopathy* has been found to significantly affect decisions made by professionals (e.g., clinicians' estimation of treatability). However, labeling children and adolescents as psychopathic does not produce more negative effects than using the term *conduct disorder* (CD). Thus, it appears that any term used to describe individuals with antisocial behavior or traits will acquire negative connotations. In fact, the perceived pejorative connotation associated with the term *callous-unemotional* (see Frick & Nigg, 2012) was one of the driving concerns behind referring to this constellation of traits as "With Limited Prosocial Emotions" in the Diagnostic and Statistical Manual of Mental Illness, Fifth Edition (DSM-5; American Psychiatric Association , 2013).

Assessing CU Traits

Several instruments are currently available to assess CU and psychopathic traits in youth. The majority originate from the Psychopathy Checklist (PCL-R; Hare, 1991, 2003) for adults, and include the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003) and the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). The Inventory of Callous-Unemotional Traits (ICU; Frick, 2004; Kimonis et al., 2008) is a derivative of the APSD and a commonly used measure of CU traits that was developed over decades to improve upon its predecessor in three important ways. First, the ICU includes a larger number of items to better capture the emotional detachment dimension of psychopathy. Second, it balances positive and negative wording of items. Third, the instrument uses a 4-point Likert-type response format to prevent against response bias and an exact middle rating. Unlike the PCL-R that has an established diagnostic cut-off score of 30 to identify adults with psychopathy (Hare, 2003), these instruments designed for youth do not have existing cut-off scores to assist practitioners in categorizing youth into those showing meaningful and clinically significant levels of CU traits.[1]

Establishing cut-off scores for measures of CU traits is of considerable importance to treatment planning for youth presenting with conduct problems. For example, several studies of justice-involved adolescents report that youth scoring high on psychopathic traits were less likely to participate in treatment, had lower quality participation, and were more likely to reoffend after treatment than those low on psychopathic traits (Gretton, Mcbride, Hare, O'Shaughnessy, & Kumka, 2001; O'Neill, Lidz, & Heilbrun, 2003; Spain, Douglas, Poythress, & Epstein, 2004). Younger samples have similarly been found to respond more poorly to traditional interventions when scoring high on CU traits (Hawes, Price, & Dadds, 2014). For example, Hawes and Dadds (2005) found that parents of boys ($M_{age} = 6.29$ years) with high CU traits were more likely to rate the discipline component (i.e., time out) of a manualized parent training program as ineffective at reducing their children's conduct problems during

treatment relative to parents whose children had oppositional defiant disorder (ODD) alone (see also Haas et al., 2011).

However, treatments tailored to the unique needs of children and adolescents with high CU traits are effective in reducing conduct problems, CU traits, and recidivism rates (Caldwell, Skeem, Salekin, & Van Rybroek, 2006; Kolko & Pardini, 2010). For example, in a study of 177 clinic-referred children, those with CU traits who received an individualized and comprehensive modular intervention evinced similar rates of improvement to other children with CD (Kolko & Pardini, 2010). These findings highlight the critical need to appropriately categorize youth with conduct problems into more homogeneous subgroups to ensure that they are receiving the optimal treatment and that time and resources are not wasted on contraindicated treatments.

The Present Study

The primary aim of the current longitudinal study was to determine cut-off scores on a parent-reported version of the ICU that optimally identified youth at risk for co-occurring stable and high CU traits and conduct problems across time. Given the difficulty in determining the optimal standard against which to develop a cut-off score, a statistical method called group-based trajectory modeling, which is widely used in clinical research (Nagin & Odgers, 2010), was applied to identify clusters of community youth presenting more homogeneous patterns of symptom presentations across childhood. ICU cut-off scores that optimally classified youth into a trajectory of high stable CU traits co-occurring with high stable conduct problems, according to mother- and father-report, were identified for the full sample and also separately by gender and age group. The second aim of the current study was to validate the ICU cut-off score using bullying as a criterion measure. This was accomplished by testing whether children scoring above the ICU cut-off scores were more likely to engage in future bullying behavior compared to children scoring below. Several studies report that youth scoring high on CU traits are more likely to engage in bullying behavior (e.g., Fanti, Frick, & Georgiou, 2009; Fanti & Kimonis, 2012, 2013; Viding, Simmonds, Petrides, & Frederickson, 2009). This link has been explained by their poor recognition of others' distress cues, their lack of concern for others' feelings, and their expectation that aggressive behavior will result in positive outcomes (Fanti & Kimonis, 2012, 2013). Bullying is a particularly appropriate outcome measure within school-age samples as it represents a developmentally appropriate form of antisocial behavior.

Method

Participants

Participants included 1,370 families in the country of Cyprus, each with a child (53.4% girls) between the ages of 6 and 12 years (M_{age} at first assessment = 9.38, SD = 1.64). Families with children in grades one through six were recruited for the study, and the resulting sample was divided evenly across grades, with approximately 16% of children in each. A cross-sequential research design combining both cross-sectional and longitudinal methods was employed. This design was selected to shorten the time required to collect longitudinal data and to approximate longitudinal growth across a

period of six years, which involved the repeated measurement (i.e., three longitudinal waves approximately six months apart) of children in the six elementary school grades. Since data were originally collected by grade cohort, the dataset was transformed to age cohorts to enable the investigation of longitudinal trajectories between the ages of 6.5 and 12 years. This design is a good approximation of a single long-term longitudinal design, and Latent Growth Models are well suited for this type of model estimation (Duncan, Duncan, & Strycker, 2013). Full information maximum likelihood fitting can be used to impute missing data across time.

Procedure

Following approval of the study by the Cyprus Ministry of Education, the second author randomly selected 26 urban and rural schools in the four school districts (Larnaka, Lemesos, Pafos, and Lefkosia) to ensure that the sample was representative of the general population. Administrators and personnel were provided a description of the study, and the boards of participating schools approved the study. Prior to data collection, signed parental consent and youth assent were obtained from 85% of participating families. Parents and children completed a battery of study questionnaires. In some cases only mother or father reports were available, although for the majority of children both parents participated in the study. At Time 1, data were collected from 1,127 mothers and 818 fathers (both parents, $n = 1,028$), at Time 2 (six months after Time 1) data were collected from 1,050 mothers and 805 fathers (both parents, $n = 991$), and at Time 3 (six months after Time 2) data were collected from 955 mothers and 732 fathers (both parents, $n = 886$). Data on bullying behavior were also collected from 719 children at Time 2 and 792 children at Time 3. Families did not receive incentives for their participation.

Measures

Conduct problems. The Checkmate plus Child Symptom Inventory for Parents-4 (CSI-4; Gadow & Sprafkin, 2002) was designed to assess symptoms of ODD (8 items; e.g., "Argues with adults") and CD (15 items; e.g., "Has stolen things from others using physical force") based on the diagnostic criteria specified in the DSM-IV (APA , 1994). This measure was administered at all three time points, and parents indicated the frequency with which their child exhibited ODD and CD symptoms on a 4-point Likert-type scale (0 = *never* to 3 = *very often).* ODD and CD subscales were combined to create an overall conduct problem score. Cronbach's α for this combined variable ranged from .86-.87 across time and depending on which parent was reporting, corresponding to good internal consistency (Barker, Pistrang, & Elliott, 1994).[2] Prior research supports the validity of parent-reported ODD and CD symptom scores from the CSI-4 in community and clinical samples in Cyprus and the U.S. (Fanti & Muñoz Centifanti, 2013; Gadow & Sprafkin, 2002).

Callous-unemotional traits. CU traits were assessed using the 24-item parent-report ICU (Frick, 2004). ICU items, such as "shows no remorse when he/she has done something wrong," were rated on a 4-point Likert-type scale (0 = *not at all* to 3 = *definitely true*), with higher scores indicating greater CU traits. The construct

validity of the ICU is supported in community, clinic-referred, and incarcerated samples of youth (e.g., Fanti, Frick, & Georgiou, 2009). For example, the ICU total score has been found to be associated with aggression, delinquency, and psychosocial and psychophysiological impairment (e.g., Fanti et al., 2009; Kimonis et al., 2008). The total score demonstrated good internal consistency (Cronbach's α ranged from .86-.88 across time). For the purposes of identifying trajectories of CU traits, mother- and father-reported ICU scores were combined in a conservative fashion by taking the higher rating between parents. This method is beneficial for circumventing potential underreporting (e.g., Pardini, Lochman, & Powell, 2007). Results of studies that combine multiple informants in this way are similar to those that have used different procedures (Piacentini, Cohen, & Cohen, 1992). To establish cut-off scores across informants, mother- and father-reported ICU scores were examined separately.

Bullying. The Student Survey of Bullying Behavior-Revised (SSBB-R; Varjas, Meyers, & Hunt, 2006) was administered at Times 2 and 3 to measure school bullying. Children self-reported whether they had engaged in different types (physical, verbal, and relational) of bullying behavior on an ordinal scale (0 = *never*, 1 = *once or twice a year*, 2 = *monthly*, 3 = *weekly*, or 4 = *daily*). The SSBB-R includes 12 items assessing bullying (e.g., "How often do you pick on younger, smaller, less powerful, or less popular kids by hitting or kicking them?") that are summed to create a total bullying scale with a range of possible scores from 0 to 48. Prior research reports that SSBB-R scores are a valid and reliable measure of school-based bullying in community samples of children and adolescents in Cyprus and the U.S. (Fanti et al., 2009; Hunt, Meyers, Jarrett, & Neel, 2005; Varjas et al., 2006). Cronbach's s for the bullying scale were 0.89 at Time 2 and 0.90 at Time 3, falling in the good to excellent range.

Plan of Analysis

We accomplished the study aims by first identifying distinct groups of developmental trajectories of conduct problems and CU traits from ages 6.5 to 12 years using Latent Class Growth Analysis (LCGA) in Mplus 6.1 for Windows (Muthén & Muthén, 2010). The purpose of this analysis was to identify children showing a trajectory of stable high scores on measures of CU traits and conduct problem symptoms from early childhood to adolescence. LCGA identifies heterogeneous latent classes based on longitudinal data by modeling the relationship between a given attribute (e.g., conduct problems or CU traits) and age. LCGA also allows for the investigation of whether change over time in the attribute is linear or whether it follows a more complex, curvilinear pattern. Full information maximum likelihood fitting was used in Mplus to retain children with incomplete assessments in the analysis. The LCGA estimation in Mplus resulted in two outputs: (a) the shape and location of the different estimated class trajectories, and (b) the posterior probability and entropy values of class membership. Average probabilities and entropy values greater than .70 indicate clear classification and greater power to predict class membership (Muthén & Muthén, 2010; Nagin, 2005).

The Bayesian Information Criterion (BIC) was one of the model fit statistics used in the current study. Models with lower BICs are preferred (Schwarz, 1978); however, since the BIC criterion tends to favor models with fewer classes by penalizing added parameters, the Lo, Mendel, Rubin (LMR) statistic was also used. Lo, Mendell, and Rubin (2001) adjusted the likelihood ratio test in order to be used in growth mixture

modeling to enable the comparison of non-nested models with different numbers of classes. The LMR statistic tests $k-1$ classes against k classes. A significant chi-square value ($p < .05$) indicates that the $k-1$ class model is rejected in favor of the k-class model. A non-significant chi square value ($p > .05$) suggests that a model with one fewer class is preferred.

Second, to identify subgroups of children with high stable CU traits co-occurring with high stable conduct problems, cross-tab analyses were performed in IBM SPSS Statistics for Windows, Version 19.0 (IBM Corp., 2010). Groups were created on the basis of membership into conduct problem and CU trajectories to identify a subgroup of children with high stable CU traits co-occurring with high stable conduct problems.

The primary purpose of the present study was to identify ICU cut-off scores for the accurate identification of the groups determined using LCGA. To accomplish this goal, intraclass correlation coefficients (ICC) were first used to test the *a priori* assumption that CU scores were stable across the three study time points. Next, sensitivity and specificity were calculated by rater using each ICU score as a potential cut-off threshold. In the context of the present study, sensitivity referred to the proportion of children in the diagnostic group of interest (e.g., high stable trajectory group) who scored at or above the cut-off threshold on the ICU, whereas specificity referred to the proportion of children not in the diagnostic group of interest who scored below the cut-off threshold (Singh, 2013a). These discriminative parameters were calculated both when "high CU/high CP" youths served as the diagnostic group of interest, as well as when the diagnostic group of interest was "high CU/high CP", "high CU/moderate CP", and "moderate CU/high CP" youths. The former analysis corresponds to Risk-Needs-Responsivity principles (Bonta & Andrews, 2006), allowing preventative resources to be allocated to only those youths at highest risk, whereas the latter represents an approach consistent with screening (Fazel, Singh, Doll, & Grann, 2012; Singh, Grann, & Fazel, 2011). Youden's (1950) J, an index for determining the cut-off threshold that balanced the cost-ratio between sensitivity and specificity, was then calculated for each score. The score with the highest J value was identified as the optimal cut-off threshold, with a cost-ratio of false positive to false negative classifications of 1:1. Cut-off thresholds were investigated for the sample overall and by gender as well as age group.[3]

Finally, independent samples t-tests were used to investigate whether children scored by parents as at or above the ICU cut-off score identified for the general sample (measured at Time 1) were more likely to self-report engaging in future bullying (Times 2 and 3) compared with children scoring below the ICU cut-off score.

Results

Descriptive Statistics

Descriptive statistics and correlations are presented in Table 1. As shown, CU traits were correlated with conduct problems across time based on both mother and father reports, and mother and father reports of CU traits and conduct problems were moderately correlated across time. Similarly, single measures and average measures ICCs were high across time between mother and father reports of CU traits (single measures ICC: T1 = .68 [.63-.71], T2 = .66 [.62-.70], T3 = .68 [.64-.72], $p < .001$; average measures ICC: T1 = .81 [.78-.83], T2 = .79 [.76-.82], T3 = .81 [.78-.84], $p < .001$) and CP (single

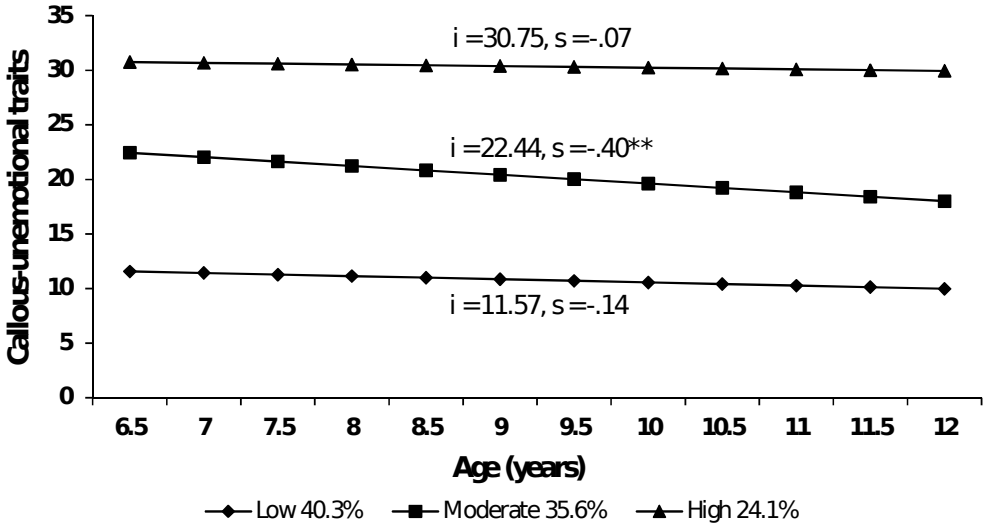

Figure 1a. LCGA model for CU traits

measures ICC: T1 = .66 [.61-.70], T2 = .66 [.62-.70], T3 = .69 [.64-.72], $p < .001$; average measures ICC: T1 = .79 [.76-.82], T2 = .80 [.77-.82], T3 = .81 [.78-.84], $p < .001$).

Identifying Trajectories of CU Traits and Conduct Problems

To identify the optimal number of trajectories for CU traits, models with one to four classes were estimated with LCGA, in accordance with the model fit indices. The BIC statistic increased from Class 3 to Class 4 and the LMR statistic fell out of significance for the four-class model, suggesting that the three-class model better fit the data (Table 2). Furthermore, the BIC for the curvilinear model (22,156.15) was higher than the BIC for the linear model (22,138.17), indicating that the linear model better represented the data. In addition, the mean probability score for the identified classes ranged from .79-.87 and the entropy value was .77, suggesting that the classes were well differentiated in the final model (Fig. 1a). Children assigned to the low risk group (40.3%) scored below average on CU traits across time. Children in the moderate risk group (35.6%) scored slightly above average on CU traits across time, although they showed decreases in these traits from early childhood to early adolescence. Children in the high stable group (24.1%) exhibited continuous high levels of CU traits, and scored approximately 1.5 standard deviations above the mean on CU traits across time. According to χ^2 analyses, boys were less likely to be in the low risk group than girls, $\chi^2(2, N = 1,192) = 7.88$, $p < .05$, $d = .25$, 95% CI [.08-.42].

To identify the optimal number of trajectories for conduct problem symptoms, models with one to four classes were estimated with LCGA. The change in the BIC statistic from class three to class four was much smaller compared to the change from

Table 1.
Descriptive statistics and correlations among the main study variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. T1 CU (father)												
2. T1 CP (father)	.46**											
3. T2 CU (father)	.56**	.35**										
4. T2 CP (father)	.34**	.69**	.42**									
5. T3 CU (father)	.59**	.38**	.59**	.38**								
6. T3 CP (father)	.38**	.63**	.37**	.71**	.47**							
7. T1 CU (mother)	.68**	.32**	.41**	.23**	.47**	.31**						
8. T1 CP (mother)	.36**	.66**	.30**	.53**	.37**	.56**	.43**					
9. T2 CU (mother)	.45**	.27**	.66**	.29**	.52**	.35**	.60**	.38**				
10. T2 CP (mother)	.28**	.53**	.35**	.66**	.34**	.57**	.31**	.69**	.45**			
11. T3 CU (mother)	.50**	.33**	.49**	.33**	.68**	.41**	.62**	.43**	.70**	.43**		
12. T3 CP (mother)	.35**	.55**	.33**	.59**	.39**	.70**	.37**	.71**	.44**	.78**	.52**	
Descriptives												
Mean	16.50	4.61	14.48	4.15	14.44	4.36	16.05	5.33	14.42	4.92	13.93	5.05
SD	9.34	4.40	9.51	4.28	9.29	4.35	9.23	4.94	9.29	4.65	9.06	5.04

Note: $^{**}p < .001$; T = Time, CU = Callous-unemotional, CP = Conduct problems.

Table 2.
Model fit statistics and posterior probabilities based on LCGA

(a)	CU traits		
Classes	BIC	Entropy	LMR
1	23,091.07	N/A	N/A
2	22,247.50	0.79	$p < .001$
3	**22138.17**	**0.77**	$\mathbf{p < .001}$
4	22141.28	0.59	$p = .11$
(b)	Conduct problems		
Classes	BIC	Entropy	LMR
1	18829.57	N/A	N/A
2	17871.78	0.89	$p < .001$
3	**17271.06**	**0.87**	$\mathbf{p < .01}$
4	17078.24	0.80	p = .08
(c)	Conduct problems		
	Low	Moderate	High
Low CU traits	0.362	0.041	0.000
Moderate CU traits	0.244	0.105	0.007
High CU traits	0.105	0.113	0.023

Note: Part a and b show model fit statistics based on the LCGA. Part c shows the proportion of children belonging to each class.

Class 2 to Class 3, which suggests that the greatest improvement in fit occurred from the two-class to the three-class model (Table 2). In addition, the LMR statistic fell out of significance for the four-class model, suggesting that the three-class model better represented the data. Moreover, the four-class model identified two very similar low classes of small theoretical importance. Accordingly, the more parsimonious three-class model was selected. The BIC (17,285.67) for the curvilinear model was higher in comparison to the BIC (17,271.06) for the linear model, indicating that the linear model better represented the data. The mean probability score for the three conduct problem classes ranged from .90-.95 and the entropy value was .87, suggesting that the classes were well separated (Fig. 1b). Children assigned to the low risk group (71.1%) exhibited low conduct problems across time. Children in the moderate risk group (25.9%) showed above average levels of conduct problems across time. Children in the high risk group (3%) showed a linear increase in conduct problems from early childhood to early adolescence, and remained at higher risk compared with low and moderate risk groups. Both the moderate and high-risk groups showed above average levels of conduct problems. According to χ^2analyses, the low group comprised more girls than boys and the moderate risk group was overrepresented by boys,

$\chi^2(2, N = 1,192) = 11.68$, $p < .01$, $d = .30$, 95% CI [.13, .49]. Cross-tab analyses were

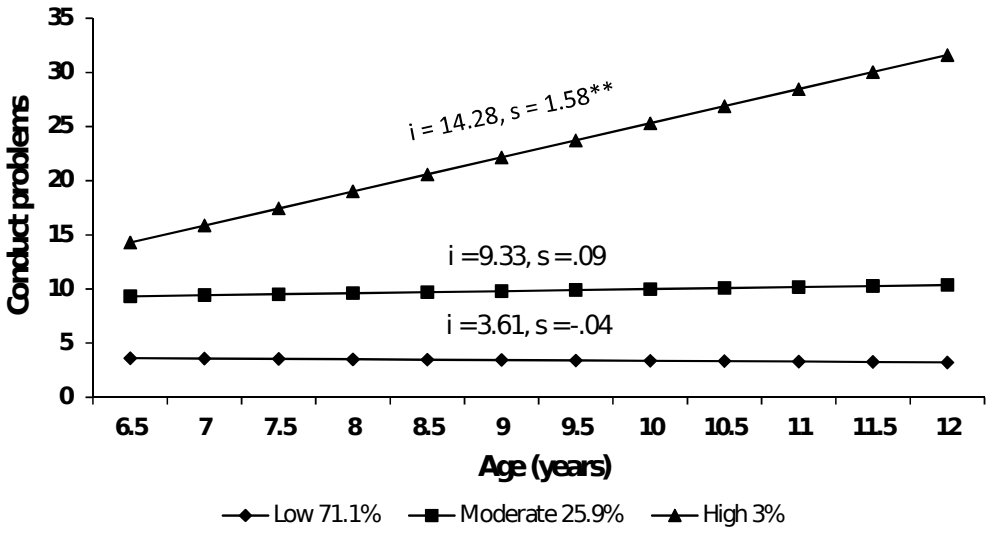

Figure 1b. LCGA model for conduct problems

used to create groups on the basis of CU trait and conduct problem trajectories. This analysis indicates the percentage of children represented in all possible joint classes between the three-class model for CU and the three-class model for conduct problems. Table 2(c) shows the proportion of children belonging to each combined group of conduct problems and CU traits. Results indicated that the majority of children in the stable high conduct problem group also fell in the stable high CU trait group (77%). Children with moderate conduct problems were more likely to score high or moderately on CU traits, and only a small percentage of children with moderate conduct problems scored low on CU traits. Finally, children with low conduct problems were more likely to score low on CU traits, although a number of children in the low conduct problems group were also classified in the moderate or high CU groups.

Identifying ICU Cut-off Scores

Prior to determining ICU cut-off scores, single measures and average measures ICCs from two-way mixed effects models for ICU scores were examined. Results indicated that resolved (i.e., combined mother and father) ICU total scores were stable across the three time points (single measures ICC = .61, $p < .001$, 95% CI [.57, .64]; average measures ICC = .82, $p < .001$, 95% CI [.80, .84]). Given these substantial to almost perfect levels of ICU score reliability (Landis & Koch, 1977), it was decided to investigate cut-off thresholds for the instrument using scores obtained at Time 1. Using the Time 1 data, sensitivity, specificity, and Youden's J values were calculated for each potential mother- and father-reported ICU score for the overall sample and separately by gender and age group. Classification into the high CU/high CP group, high CP/moderate CU group, or

moderate CP/high CU group were used as the diagnostic outcome. The cut-off scores that optimized both sensitivity and specificity on the ICU are presented in Table 3 for mother-report and Table 4 for father-report (full statistical details available from the authors). Overall, for mother-report for the full sample and for boys alone an ICU total score of 24 best identified those youth in the high CU/high CP trajectory, whereas a higher threshold score of 27 was identified for girls. For father-report for the full sample and for girls alone this score was 27, but was lower for boys at 25. Across raters, the range of cut-off scores varied by age and ranged from 16 to 29. Tables 3 and 4 also present cut-off scores for optimally predicting alternative combinations of CU and CP levels.

Criterion Validity of the Identified Cut-off Score

On average, relative to children rated by parents as below the general sample ICU cut-off score (total score = 24 for mother-report and 27 for father-report), those scoring at or above it were more likely to engage in bullying at Time 2, according to both mother-reported ($M = 2.86$, $SD = 5.75$ versus $M = 5.35$, $SD = 8.59$, $d = .39$, 95% CI [.23, .55], $t(659) = 3.25$, $p < .001$) and father-reported CU traits ($M = 3.03$, $SD = 6.01$ versus $M = 5.66$, $SD = 8.82$, $d = .40$, 95% CI [.20, .60], $t(534) = 4.12$, $p < .001$). This finding was consistent at Time 3 for mother-reported ($M = 4.14$, $SD = 6.86$ versus $M = 6.21$, $SD = 8.02$, $d = .29$, 95% CI [.10, .55], $t(599) = 2.81$, $p < .01$) and father-reported ($M = 4.41$, $SD = 7.01$ versus $M = 6.01$, $SD = 8.05$, $d = .22$, 95% CI [.08, .42], $t(490) = 2.20$, $p < .05$) CU traits.

Discussion

The present study aimed to establish cut-off scores for the parent-report version of a commonly used comprehensive measure of CU traits in youth—the Inventory of Callous-Unemotional Traits. This was accomplished by first identifying trajectories of youth with more homogeneous patterns of conduct problems and CU traits across childhood. This analysis revealed that 2.3% of youth in this nationally representative Cypriot community sample showed high stable conduct problem symptoms co-occurring with high stable CU traits. Findings indicated that using cut-off scores of 24 and 27 on the mother- and father-reported ICU (respectively) most accurately identified this group of children. Consistent with prior research, CU traits were moderately to highly stable across the three study time points (Frick, Kimonis, Dandreaux, & Farell, 2003). Importantly, the identified cut-off scores also predicted future child-reported bullying behavior, an established external correlate to CU (Fanti & Kimonis, 2013).

Currently there is no clear consensus regarding the most appropriate base rate for defining non-normative and impairing levels of CU traits and this base rate will likely depend on the setting (e.g., community, clinic-referred, incarcerated), number and type of informants, and the purpose for making this diagnosis (e.g., the importance of avoiding false positives versus avoiding false negatives; Kahn, Frick, Youngstrom, Findling, & Youngstrom, 2012). However, Frick and Viding (2009) estimated that 2–4% of all children show a joint CP+CU presentation on the basis of CU prevalence estimates obtained within samples of children with conduct problems, and in their Fast

Table 3.
Summary of cut-off thresholds using mother-reported CU scores ($n = 1,208$)

Analysis	Cut-off 1	Sensitivity	Specificity	Youden's J	Cut-off 2	Sensitivity	Specificity	Youden's J
Overall	24	0.69	0.78	0.47	17	0.86	0.61	0.48
Boys	24	0.86	0.76	0.62	16	0.91	0.59	0.5
Girls	27	0.50	0.86	0.36	17	0.84	0.61	0.45
Age 7	16	0.88	0.57	0.44	16	0.94	0.63	0.57
Age 8	29	0.75	0.89	0.64	26	0.54	0.86	0.40
Age 9	16	0.96	0.63	0.60	28	1.00	0.85	0.85
Age 10	21	1.00	0.71	0.71	17	0.93	0.61	0.54
Age 11	13*	1.00	0.45	0.45	22	0.75	0.80	0.55
Age 12	26	1.00	0.75	0.75	20	0.70	0.64	0.34

Note: Cut-off 1 is for identifying youth on trajectories of high stable co-occurring conduct problems and CU traits; Cut-off 2 is for identifying youth on trajectories of either high stable co-occurring conduct problems and CU traits, high stable conduct problems-moderate CU traits, or moderate conduct problems-high stable CU traits; *The low base rate of membership in the high CU-conduct problems group for this age group suggests that this cut-off score is a statistical artifact.

Table 4.
Summary of cut-off thresholds using father-reported CU scores ($n = 818$)

Analysis	Cut-off 1	Sensitivity	Specificity	Youden's J	Cut-off 2	Sensitivity	Specificity	Youden's J
Overall	27	0.70	0.83	0.53	22	0.82	0.76	0.58
Boys	25	0.80	0.73	0.53	22	0.93	0.74	0.67
Girls	27	0.70	0.87	0.57	22	0.69	0.77	0.47
Age 7	25	0.83	0.80	0.63	22	0.91	0.77	0.68
Age 8	35	0.33	0.99	0.33	23	0.82	0.78	0.60
Age 9	28	1.00	0.87	0.87	21	0.84	0.75	0.59
Age 10	22	1.00	0.69	0.69	20	0.83	0.72	0.54
Age 11	37*	1.00	0.99	0.99	24	0.88	0.83	0.71
Age 12	27	1.00	0.75	0.75	25	0.88	0.80	0.68

Note: Cut-off 1 is for identifying youth on trajectories of high stable co-occurring conduct problems and CU traits; Cut-off 2 is for identifying youth on trajectories of either high stable co-occurring conduct problems and CU traits, high stable conduct problems-moderate CU traits, or moderate conduct problems-high stable CU traits; *The low base rate of membership in the high CU-conduct problems group for this age group suggests that this cut-off score is a statistical artifact.

Track sample McMahon et al. (2010) reported a prevalence rate of 1.2%. These estimates are within range of the 2.3% prevalence rate for youth with high stable co-occurring CU traits and CP in the present study and lend confidence to the generalizability of the findings.

Whereas the 3% prevalence of conduct problems was consistent with large epidemiological studies, the prevalence of high stable CU traits was relatively high in this sample of Cypriot youth, whether presenting alone (24%) or within youth with co-occurring high stable conduct problems (77%). In comparison, prior studies of children in the U.K. that used items from non-standardized instruments report prevalence estimates between 3-5% for high CU traits in the absence of conduct problems (Fontaine, McCrory, Boivin, Moffitt, & Viding, 2011; Rowe et al., 2010). Average ICU total scores in the present study (father-report: $M = 16.50$, $SD = 9.34$, mother-report: $M = 16.05$, $SD = 9.23$) contradict the possibility that Cypriot youth are on average more callous and unemotional than youth from the U.K., for whom mean ICU total scores were almost one standard deviation higher in a large ($N = 704$) community sample of 11-13 year-old children ($M = 24.72$, $SD = 9.01$; Viding, Simmonds, Petrides, & Frederickson, 2009). Future research is necessary to understand why a quarter of this sample presented with ICU scores approximately 1.5 standard deviations above the sample mean.

Children presenting with CU traits without conduct problems have been relatively understudied in the literature, likely due in large part to their lesser impairment relative to those with co-occurring conduct problems. For example, children high on CU traits alone are at lower risk for adverse outcomes compared to children high on both dimensions (Fanti, 2013; Rowe et al., 2010). Focusing on children high on CU traits alone in future research may help elucidate those factors that protect at-risk children from maladaptive outcomes and a destructive course of antisocial and aggressive behavior. Perhaps importantly, although a high percentage of children were identified with high CU traits, only a small minority of children scored consistently high on both conduct problems and CU traits. According to a recent review, it appears to be the combination of conduct problems and CU traits, and not CU traits alone, that leads to future maladaptive outcomes, representing a more severely impaired group of youth (Frick et al., 2014).

Extant research suggests that children with conduct problems and CU traits differ in important ways from children with conduct problems alone (Frick et al., 2014). For example, their antisocial behaviors tend to be more severe and aggressive, earlier starting and more stable than children without CU traits. To illustrate the utility of the identified cut-off scores in identifying children consistent with this profile, based on these data, a score of 27 or higher on the ICU (father-report) correctly identified 70% of children within the stable high CU/high CP trajectory whereas a score below 27 correctly identified 83% of children not in this high-risk group. When scored at or above this ICU threshold at Time 1, children were significantly more likely to report engaging in bullying behaviors six months and one year later, than children scoring below the cut-off score. This finding is consistent with cross-sectional (Fanti et al., 2009; Viding et al., 2009) and longitudinal research (Fanti & Kimonis, 2012) reporting an association between bullying and CU traits. It also provides support for the predictive validity of the identified ICU cut-off score with respect to its ability to predict an external correlate to CU traits measured from a different reporter.

Relating the identified cut-off scores to the assessment of CU traits in previous research, the cut-off scores are roughly half a standard deviation above the mean self-reported ICU total score ($M = 22.50$, $SD = 8.20$) reported for a mixed sample ($N = 383$) of European community, clinic-referred, detained, and child welfare-involved boys (8–20 years; Feilhauer, Cima, & Arntz, 2012), and (for father-report) roughly equivalent to the mean parent-reported ICU total score ($M = 27.19$, $SD = 12.20$) reported for a sample ($N = 94$) of American adolescent boys (12–18 years) adjudicated of a sexual offense (White, Cruise, & Frick, 2009), recognizing important differences in age, informant, and setting across studies. When the intention is to screen youth for the purposes of preventive intervention, scores ≥ 17 for mother-report and ≥ 22 for father-report on the ICU most accurately identified those in the overall sample who fell in the groups showing moderate trajectories of CU or conduct problems combined with high stable trajectories of the other. Using either method, cut-off scores varied somewhat when examined separately by gender or age group. Whereas the cut-off scores generated in this study may have utility for identifying youth at risk for severe and stable conduct problems co-occurring with CU traits to deliver nuanced interventions, it is important to note that many research questions may be better addressed using continuous measures of CU traits that utilize the complete information provided by the Likert-type response format of most CU measures (Thornton, Frick, Crapanzano, & Terranova, 2013).

The results of the present study must be considered within the context of its relative strengths and limitations. First, although both mother- and father-reports of CU traits were examined, data on the presence of CU traits in the school setting according to teacher reports was not available. Diagnostic systems, such as the DSM-5 (APA , 2013) explicitly recognize the importance of carefully considering multiple sources of information including self- and informant report (e.g., parents, teachers, peers, other family members) from sources who have known the child for a significant period of time when evaluating the presence of non-normative CU traits. Second, symptoms of ODD and CD were combined in order to identify children with conduct problems more generally who are at greatest risk for continuing and severe impairment across childhood. Thus, these data cannot speak to the utility of the identified cut-off scores within the context of the "With Limited Prosocial Emotions" CU specifier to the diagnosis of CD.

Third, bullying is a conduct disorder symptom (i.e., often bullies, threatens, or intimidates others), which lends to the possibility of criterion contamination when using it as an external criterion measure. However, different sources rated measures of conduct problems and bullying, lending some confidence to the ability of parent-reported ICU cut-off scores to predict child-reported bullying. While using self-report to assess bullying runs the risk of response bias given links between psychopathy and deception (see Rogers & Cruise, 2000), some research suggests that antisocial attitudes and behavior are more reliably assessed using self-report than caregiver-report, especially into adolescence (Jolliffe et al., 2003). Future research might consider measuring bullying using a multi-method approach (self-, parent-, teacher-report), and incorporating other external correlates that are important to the construct of CU traits, such as cognitive (e.g., punishment insensitivity), emotional (e.g., distress insensitivity) and biological (e.g., reduced amygdala activation) variables (Frick et al., 2014). Fourth, the sample comprised Greek-speaking children from the

Republic of Cyprus and findings must be replicated in samples from other countries and with adolescent populations to determine the predictive validity of identified cut-off scores with respect to a variety of outcomes to determine their generalizability. Finally, cut-off scores could be established for all age groups except 11-year-olds, for whom there was a very low base rate of membership into the high CU/high CP trajectory group. Thus, future research is needed to establish whether similar cut-off scores identify at-risk children in this age group relative to other ages.

Important strengths of this study were the large sample size, collection of longitudinal data, and the use of both mother and father-reported CU traits and conduct problems to address the possibility of under-reporting of symptoms. We also employed advanced statistical modeling, adopted from the field of criminology and becoming increasingly popular in clinical research (Nagin & Odgers, 2010), to identify a subset of youth showing high chronic CU and conduct problem symptoms in order to optimize prediction from possible ICU cut-off scores. In the context of these limitations and strengths, our results provide some useful information for defining significant levels of CU traits among community-based samples of school-aged youth.

The Risk-Needs-Responsivity principles suggest that preventative resources should be allocated to only those youths at highest risk (Bonta & Andrews, 2006). The process of identifying which community-based youths are at greatest risk for future impairment to administer empirically supported interventions to them may be assisted by applying the ICU cut-off scores identified in the present study. Utilizing Youden's J statistic to identify optimal scores assumes that the costs of false positives and false negatives are equally deleterious, meaning that decision-making using the instrument will be neither over-inclusive nor under-inclusive. There were two primary reasons for selecting this approach. First, from a diagnostic perspective, youth (inappropriately) assigned the CU specifier (i.e., false positives) would be at no greater risk from suffering from this label as they would from the CD label that must be diagnosed first according to the DSM-5 (see Murrie, Boccaccini, McCoy, & Cornell, 2007). Second, failing to classify a youth with CD with the "With Limited Prosocial Emotions" specifier (i.e., false negative) may result in delivery of parent-training interventions that are the gold-standard for children with conduct problems but are less effective for those with CU traits (Hawes et al., 2014). This may delay CP+CU children from receiving appropriate treatment (i.e., intensive interventions tailored to their unique emotional and cognitive characteristics) that have been found to improve their CU traits, conduct problems, and other antisocial outcomes (Kolko & Pardini, 2010).

International surveys of practitioners who use structured instruments in psychological assessments have clearly shown a preference for categorical methods of risk communication (Singh, 2013b). However, as recent research has suggested that antisociality may be a taxometrically continuous rather than dichotomous construct, some researchers have argued that cut-off scores should not be established for personality measures such as the ICU (Marcus, Lilienfeld, Edens, & Poythress, 2006). It may be that the ICU could be used most efficiently following a two-stepped decision-making strategy. First, the cut-off scores identified in the current study would be used to determine whether resources should be allocated in a given case. Second, risk ratio and percentile information derived from continuous ICU scores would be used to determine how many resources should be allocated and the nature of those resources. Future research could examine whether risk ratios and percentiles from large-scale or

jurisdiction-specific normative studies result in the most effective risk management plans. In conclusion, this promising line of research provides a step towards appropriately sorting antisocial youth into more homogenous subgroups to improve outcomes for a particularly impaired population.

References

American Psychiatric Association . (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association . (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Barker, C., Pistrang, N., & Elliott, R. (1994). *Research methods in clinical and counselling psychology.* Chichester, UK: Wiley.

Bonta, J., & Andrews, D. A. (2006). *Risk-need-responsivity model for offender assessment and rehabilitation.* Ottawa, ON: Public Safety and Emergency Preparedness. (Cat.No. PS3-1/2007-6)

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*(4), 371–380. doi: 10.1111/j.1469-7610.2011.02504.x

Caldwell, M., Skeem, J. L., Salekin, R. T., & Van Rybroek, G. (2006). Treatment response of adolescent offenders with psychopathy features: A 2-year follow-up. *Criminal Justice and Behavior*, *33*(5), 571–596. doi: 10.1177/0093854806288176

Campbell, M. A., Porter, S., & Santor, D. (2004). Psychopathic traits in adolescent offenders: an evaluation of criminal history, clinical, and psychosocial correlates. *Behavioral Sciences & the Law*, *22*(1), 23–47. doi: 10.1002/bsl.572

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2013). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Routledge Academic.

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, J., Norman G. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, *115*(1), 131–144. doi: 10.1037/0021-843x.115.1.131

Fanti, K. A. (2013). Individual, social, and behavioral factors associated with co-occurring conduct problems and callous-unemotional traits. *Journal of Abnormal Child Psychology*, *41*(5), 811–824. doi: 10.1007/s10802-013-9726-z

Fanti, K. A., Frick, P. J., & Georgiou, S. (2009). Linking callous-unemotional traits to instrumental and non-instrumental forms of aggression. *Journal of Psychopathology and Behavioral Assessment*, *31*(4), 285–298. doi: 10.1007/s10862-008-9111-3

Fanti, K. A., & Kimonis, E. R. (2012). Bullying and victimization: The role of conduct problems and psychopathic traits. *Journal of Research on Adolescence*, *22*(4), 617–631. doi: 10.1111/j.1532-7795.2012.00809.x

Fanti, K. A., & Kimonis, E. R. (2013). Dimensions of juvenile psychopathy distinguish "bullies," "bully-victims," and "victims". *Psychology of Violence*, *3*(4), 396–409. doi: 10.1037/a0033951

Fanti, K. A., & Muñoz Centifanti, L. C. (2013). Childhood callous-unemotional traits moderate the relation between parenting distress and conduct problems over time.

Child Psychiatry and Human Development, *45*(2), 173–184. doi: 10.1007/s10578-013-0389-3

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). The prediction of violence and antisocial behaviour: A systematic review and meta-analysis of the utility of risk assessment instruments in 73 samples involving 24,827 individuals. *British Medical Journal*, *345*, e4692. doi: 10.1136/bmj.e4692

Feilhauer, J., Cima, M., & Arntz, A. (2012). Assessing callousunemotional traits across different groups of youths: Further cross-cultural validation of the inventory of callousunemotional traits. *International Journal of Law and Psychiatry*, *35*(4), 251–262. doi: 10.1016/j.ijlp.2012.04.002

Fontaine, N. M. G., McCrory, E. J. P., Boivin, M., Moffitt, T. E., & Viding, E. (2011). Predictors and outcomes of joint trajectories of callousunemotional traits and conduct problems in childhood. *Journal of Abnormal Psychology*, *120*(3), 730–742. doi: 10.1037/a0022620

Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The Psychopathy Checklist: Youth Version manual*. Toronto: Multi-Health Systems.

Frick, P. J. (2004). *The inventory of callous-unemotional traits*. New Orleans: The University of New Orleans. Unpublished rating scale.

Frick, P. J., & Hare, R. D. (2001). *Antisocial process screening device (APSD)*. Toronto: Multi-health systems.

Frick, P. J., Kimonis, E. R., Dandreaux, D. M., & Farell, J. M. (2003). The four-year stability of psychopathic traits in non-referred youth. *Behavioral Sciences & the Law*, *21*(6), 713–736. doi: 10.1002/bsl.568

Frick, P. J., & Nigg, J. T. (2012). Current issues in the diagnosis of attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. *Annual Review of Clinical Psychology*, *8*(1), 77–107. doi: 10.1146/annurev-clinpsy-032511-143150

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), 1–57. doi: 10.1037/a0033076

Frick, P. J., & Viding, E. (2009). Antisocial behavior from a developmental psychopathology perspective. *Development and Psychopathology*, *21*(4), 1111–1131. doi: 10.1017/s0954579409990071

Gadow, K. D., & Sprafkin, J. (2002). *Child Symptom Inventory–4 screening and norms manual*. Stony Brook, NY: Checkmate Plus.

Gretton, H. M., Mcbride, M., Hare, R. D., O'Shaughnessy, R., & Kumka, G. (2001). Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior*, *28*(4), 427–449. doi: 10.1177/009385480102800403

Haas, S. M., Waschbusch, D. A., Pelham, W. E., King, S., Andrade, B. F., & Carrey, N. J. (2011). Treatment response in CP/ADHD children with callous/unemotional traits. *Journal of Abnormal Child Psychology*, *39*(4), 541–552. doi: 10.1007/s10802-010-9480-4

Hare, R. D. (1991). *The psychopathy checklist-revised* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

Hare, R. D. (2003). *The psychopathy checklist-revised* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*, *62*(2), 387–397. doi: 10.1037/0022-006x.62.2.387

Hawes, D. J., & Dadds, M. R. (2005). The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting and Clinical Psychology*, *73*(4), 737741. doi: 10.1037/0022-006x.73.4.737

Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical Child and Family Psychology Review*, *17*(3), 248–267. doi: 10.1007/s10567-014-0167-1

Hunt, M. H., Meyers, J., Jarrett, O., & Neel, J. (2005). *Student survey of bullying behavior: Preliminary development and results from six elementary schools.* Atlanta, Georgia: Georgia State University Center for Research on School Safety, School Climate and Classroom Management.

IBM Corp. (2010). *IBM SPSS statistics for Windows, Version 19.0.* Armonk, NY: IBM Corp.

Jolliffe, D., Farrington, D. P., Hawkins, J. D., Catalano, R. F., Hill, K. G., & Kosterman, R. (2003). Predictive, concurrent, prospective and retrospective validity of self-reported delinquency. *Criminal Behaviour and Mental Health*, *13*(3), 179197. doi: 10.1002/cbm.541

Kahn, R. E., Frick, P. J., Youngstrom, E., Findling, R. L., & Youngstrom, J. K. (2012). The effects of including a callous-unemotional specifier for the diagnosis of conduct disorder. *Journal of Child Psychology and Psychiatry*, *53*(3), 271–282. doi: 10.1111/j.1469-7610.2011.02463.x

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., ... Morris, A. S. (2008). Assessing callousunemotional traits in adolescent offenders: Validation of the Inventory of Callous–Unemotional Traits. *International Journal of Law and Psychiatry*, *31*(3), 241–252. doi: 10.1016/j.ijlp.2008.04.002

Kolko, D. J., & Pardini, D. A. (2010). ODD dimensions, ADHD, and callousunemotional traits as predictors of treatment response in children with disruptive behavior disorders. *Journal of Abnormal Psychology*, *119*(4), 713–725. doi: 10.1037/a0020910

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lee, Z., Klaver, J. R., Hart, S. D., Moretti, M. M., & Douglas, K. S. (2009). Short-term stability of psychopathic traits in adolescent offenders. *Journal of Clinical Child & Adolescent Psychology*, *38*(5), 595–605. doi: 10.1080/15374410903103536

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review*, *14*(1), 17–38. doi: http://dx.doi.org/10.1016/0272-7358(94)90046-9

Lo, Y., Mendell, N., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. doi: 10.1093/biomet/88.3.767

Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, *116*(1), 155–165. doi: 10.1037/0021-843x.116.1.155

Marcus, D. K., Lilienfeld, S. O., Edens, J. F., & Poythress, N. G. (2006). Is antisocial

personality disorder continuous or categorical? a taxometric analysis. *Psychological Medicine*, *36*(11), 1571–1581. doi: 10.1017/s0033291706008245

Mash, E., & Dozois, D. (2003). Child psychopathology: A developmental-systems perspective. In E. Mash & R. A. Barkley (Eds.), *Child psychopathology* (2nd ed., pp. 3–71). New York: Gilford Press.

McMahon, R. J., Witkiewitz, K., & Kotler, J. S. (2010). Predictive validity of callousunemotional traits measured in early adolescence with respect to multiple antisocial outcomes. *Journal of Abnormal Psychology*, *119*(4), 752–763. doi: 10.1037/a0020796

Murrie, D. C., Boccaccini, M. T., McCoy, W., & Cornell, D. G. (2007). Diagnostic labeling in juvenile court: How do descriptions of psychopathy and conduct disorder influence judges? *Journal of Clinical Child & Adolescent Psychology*, *36*(2), 228–241. doi: 10.1080/15374410701279602

Murrie, D. C., Marcus, D. K., Douglas, K. S., Lee, Z., Salekin, R. T., & Vincent, G. (2007). Youth with psychopathy features are not a discrete class: a taxometric analysis. *Journal of Child Psychology and Psychiatry*, *48*(7), 714–723. doi: 10.1111/j.1469-7610.2007.01734.x

Muthén, L., & Muthén, B. (2010). *Mplus User's Guide* (6th ed.). Los Angeles: Muthén & Muthén.

Nagin, D. S. (2005). *Group-based modeling of development.* Boston, MA: Harvard University Press.

Nagin, D. S., & Odgers, C. L. (2010). Group-based trajectory modeling in clinical research. In Nolen-Hoekland, T. S. Cannon, & T. Widger (Eds.), *Annual review of clinical psychology.* Palo Alto, CA: Annual Reviews.

O'Neill, M. L., Lidz, V., & Heilbrun, K. (2003). Adolescents with psychopathic characteristics in a substance abusing cohort: Treatment process and outcomes. *Law and Human Behavior*, *27*(3), 299–313. doi: 10.1023/a:1023435924569

Pardini, D. A., Lochman, J. E., & Powell, N. (2007). The development of callous-unemotional traits and antisocial behavior in children: are there shared and/or unique predictors? *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 319–333. doi: 10.1080/15374410701444215

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology*, *20*(1), 51–63. doi: 10.1007/bf00927116

Rogers, R., & Cruise, K. (2000). Malingering and deception among psychopaths. In C. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 269–284). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

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*(6), 688–695. doi: 10.1111/j.1469-7610.2009.02199.x

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Seagrave, D., & Grisso, T. (2002). Adolescent development and the measurement of juvenile psychopathy. *Law and Human Behavior*, *26*(2), 219–239. doi: 10.1023/a:

1014696110850

Singh, J. P. (2013a). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law*, *31*(1), 8–22. doi: 10.1002/bsl.2052

Singh, J. P. (2013b). *Violence risk assessment in adults and juveniles: The state of the art.* (Speech presented at Harvard University, Cambridge, MA, USA)

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*(3), 499–513. doi: 10.1016/j.cpr.2010.11.009

Spain, S. E., Douglas, K. S., Poythress, N. G., & Epstein, M. (2004). The relationship between psychopathic features, violence and treatment outcome: the comparison of three youth measures of psychopathic features. *Behavioral Sciences & the Law*, *22*(1), 85–102. doi: 10.1002/bsl.576

Thornton, L. C., Frick, P. J., Crapanzano, A. M., & Terranova, A. M. (2013). The incremental utility of callous-unemotional traits and conduct problems in predicting aggression and bullying in a community sample of boys and girls. *Psychol Assess*, *25*(2), 366–378. doi: 10.1037/a0031153

Varjas, K., Meyers, J., & Hunt, M. H. (2006). *Student survey of bullying behaviorRevised 2 (SSBB-R2).* Atlanta, Georgia: Georgia State University, Center for Research on School Safety, School Climate and Classroom Management.

Vasey, M. W., Kotov, R., Frick, P. J., & Loney, B. R. (2005). The latent structure of psychopathy in youth: A taxometric investigation. *Journal of Abnormal Child Psychology*, *33*(4), 411–429.

Viding, E., Simmonds, E., Petrides, K. V., & Frederickson, N. (2009). The contribution of callous-unemotional traits and conduct problems to bullying in early adolescence. *Journal of Child Psychology and Psychiatry*, *50*(4), 471–481. doi: 10.1111/j.1469-7610.2008.02012.x

White, S. F., Cruise, K. R., & Frick, P. J. (2009). Differential correlates to self-report and parent-report of callous–unemotional traits in a sample of juvenile sexual offenders. *Behavioral Sciences & the Law*, *27*(6), 910–928. doi: 10.1002/bsl.911

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35.

Footnotes

[1] Although no formal cut-off score has been identified for the PCL:YV, some studies have used scores of 25 or 30 to designate youth as psychopathic (e.g., Campbell, Porter, & Santor, 2004; Gretton, Mcbride, Hare, O'Shaughnessy, & Kumka, 2001).

[2] Reliability coefficients $< .60$ are considered insufficient, .60–.69 marginal, .70–.79 acceptable, .80–.89 good, and $\geq .90$ excellent.

[3] Due to small sample sizes, cut-off thresholds were not calculated for ages 6 and 13 year.

Received: Sept. 17, 2014
Revision Received: Oct. 25, 2014
Accepted: Oct. 30, 2014