Assessing Callous-Unemotional Traits in Adolescents: Validity of the Inventory of Callous-Unemotional Traits Across Gender, Age, and Community/Institutionalized Status

S. Pihet · S. Etter · M. Schmid · E. R. Kimonis

Abstract Callous-unemotional (CU) traits identify adolescents at high risk for severe and recurrent antisocial behaviour and are included in the DSM-5 as a specifier to conduct disorder. The Inventory of Callous-Unemotional Traits (ICU) might be the most cost-effective screening instrument for detecting CU traits in youth. We aimed to address an important gap in research by comparing the factor structure and psychometric properties of ICU scores across gender, age, and setting (community versus institutionalized). Community-based ($n=397$) and institutionalized ($n=164$) adolescent boys and girls completed self-reported measures of psychopathic traits (including the ICU), externalizing problems and aggression, and a laboratory measure of emotional recognition; the staff reported antisocial behavior for institutionalized children. Factor-analytic results indicated that a three-factor bifactor model best fit the data across samples, with measurement invariance across gender, age, and setting, supporting the construct validity of the ICU. In support of its criterion validity, across groups ICU scores were positively correlated with analogous dimensions from the Youth Psychopathic Traits Inventory, measures of proactive aggression, and self- and staff-reported antisocial behavior, and were associated with poorer accuracy in recognizing distress-related emotions. Our findings thus support the overall utility of the ICU for assessing CU traits in adolescent populations regardless of gender, age, and community or institutionalized status, but suggest a need for refining its assessment of the shallow affect dimension.

Keywords Callous-unemotional traits · Adolescence · Community · Conduct disorder · Psychometrics

Adolescents with callous-unemotional (CU) traits tend to lack empathy, and are often perceived by others as cold-hearted and uncaring about others' feelings and their own performance. Highly focused on their current personal goals, they seem to be efficient at using others to achieve them. In addition, hurting others or breaking rules to achieve their goals seems to result in little experience of guilt and remorse (Frick and White 2008). Longitudinal research suggests that CU traits may be a developmental precursor to psychopathic traits in adulthood (e.g. Lynam et al. 2007). Similar to how psychopathic traits distinguish a unique subgroup of adult offenders, CU traits identify antisocial adolescents who show more severe and recurrent antisocial behaviour, as well as poor response to traditional treatments (e.g. Frick and Dickens 2006). This ability to identify a high-risk subgroup of delinquent adolescents has contributed to the inclusion of CU traits in the fifth revision of the *Diagnostic and Statistical Manual* (DSM-5; APA, 2013) as a specifier (called "with limited prosocial emotions") to conduct disorder (CD). Potentially, large numbers of clinicians and researchers will thus need valid cost-effective tools to screen CU traits among diverse samples of youth.

The 24-item Inventory of Callous-Unemotional Traits (ICU) was systematically developed over two decades, providing one of the most comprehensive yet brief measures of

S. Pihet (✉)
School of Health Sciences Fribourg, University of Applied Sciences and Arts Western Switzerland, Route des Cliniques 15, 1700 Fribourg, Switzerland
e-mail: Sandrine.Pihet@hefr.ch

S. Pihet · S. Etter
Department of Psychology, University of Fribourg, Fribourg, Switzerland

M. Schmid
Department of Child and Adolescent Psychiatry, University of Basel, Basel, Switzerland

E. R. Kimonis
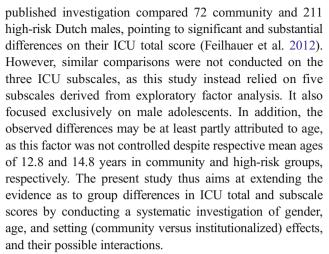School of Psychology, University of New South Wales, Sydney, Australia

CU traits currently available. Responses are provided on a 4-point Likert scale (0=*not at all true*, 1=*somewhat true*, 2=*very true*, and 3=*definitely true*). Factor analytic research identified three factors (Essau et al. 2006): *Callousness* (11 items about lack of guilt and empathy); *Uncaring* (8 items measuring the lack of care about ones' performance and the well-being of others); and *Unemotional* (5 items about restricted emotional expression). Initial validation studies suggested that the ICU was a promising instrument for assessing CU traits in non-institutionalized adolescents, although its Unemotional scale suffered from low internal consistency and sometimes poor criterion validity (Essau et al. 2006; Ciucci et al. 2014; Feilhauer et al. 2012; Kimonis et al. 2008b; Fanti et al. 2009; Roose et al. 2010). A number of issues still require investigation: 1) gender and age differences have scarcely been tested, and differences between community and institutionalized adolescents have only been investigated once in a male-only sample heterogeneous in age; 2) the stability of the ICU factor structure, or its measurement invariance across samples, has only been tested in two studies among community adolescents; and 3) confirming the stability of the ICU criterion validity across samples and using a multi-method approach would further support its psychometric soundness.

Group Differences

Correlational studies have shown that CU traits are fairly stable from late childhood into early adolescence (e.g. Muñoz and Frick 2007) and from adolescence to adulthood (e.g. Blonigen et al. 2006). However, correlations do not provide information about changes in the absolute level of the sample mean, and little is currently known about age differences in CU traits, requiring further investigation. Two ICU validation studies documented significantly higher ICU scores in 15–16 year-olds compared to 13–14 and 17–18 year-olds (Essau et al. 2006), and in 8th graders compared to 6th graders (Ciucci et al. 2014), while no age differences were found in detained male adolescents (White et al. 2009). Taken together, these findings suggest a need for age-specific reference values, although additional research with institutionalized samples is needed.

While most studies converge in showing that boys are at higher risk for developing aggressive problems than girls (e.g. Card et al. 2008), only four prior studies assessed gender differences in ICU scores in community samples (Essau et al. 2006; Decuyper et al. 2011; Ciucci et al. 2014; Fanti et al. 2009), and one in adjudicated youth (Stickle et al. 2012). They all observed significantly higher ICU scores among boys relative to girls.

Currently, most prior studies have focused either on community or on institutionalized samples. Only one recently

published investigation compared 72 community and 211 high-risk Dutch males, pointing to significant and substantial differences on their ICU total score (Feilhauer et al. 2012). However, similar comparisons were not conducted on the three ICU subscales, as this study instead relied on five subscales derived from exploratory factor analysis. It also focused exclusively on male adolescents. In addition, the observed differences may be at least partly attributed to age, as this factor was not controlled despite respective mean ages of 12.8 and 14.8 years in community and high-risk groups, respectively. The present study thus aims at extending the evidence as to group differences in ICU total and subscale scores by conducting a systematic investigation of gender, age, and setting (community versus institutionalized) effects, and their possible interactions.

Another important question is whether specific reference values are needed for different language versions and/or cultural contexts (LV/CC). The 10 previous validation studies summarized in a recent meta-analysis (Feilhauer et al. 2012) used four different LV to gather data in six CC. Comparing the ICU total scores from six community and four detained samples of adolescents, the authors concluded that "the ICU does not differentiate between detained and community samples" (p. 260). We consider this conclusion premature as the community samples were assessed using the Dutch (in the Nederland; Feilhauer et al. 2012; in Belgium; Roose et al. 2010), English (in the UK; Viding et al. 2009; in the US; Muñoz et al. 2011), German (Essau et al. 2006), and Greek (in Cyprus; Fanti et al. 2009) versions of the ICU, while the detained samples were assessed using the English (in the US; Kimonis et al. 2008b; White et al. 2009; Muñoz et al. 2008) or Dutch (in the Netherland; Feilhauer et al. 2012) versions. Further research should thus disentangle LV/CC effects from other group differences. The present study therefore firstly aimed at investigating group differences within one LV/CC, and secondly the differences between the reference values of the present study and those of prior ICU validation studies using other LV in samples similar in gender, age, and setting.

Factor Structure of the ICU

Six studies (see Table 3) have tested the factor structure of the ICU in community or institutionalized adolescents with Confirmatory Factor Analysis (CFA). They compared models comprising 1) one general factor, 2) three correlated factors (Callous, Unemotional, and Uncaring), and 3) three specific plus one general factor (bifactor model). They all confirmed the superiority of the bifactor model, i.e. a model with a general CU factor underlying each of the items along with three independent factors (Callous, Unemotional, and Uncaring) each accounting for the unique variance in their

respective set of items, over and above the general factor (for a detailed description of this bifactor model, as well as a discussion of its advantages over the hierarchical model, see Kimonis et al. 2008b). One study suggested that the factor structure might be improved by removing two items with poor loadings (Kimonis et al. 2008b).

Measurement invariance tests whether the same construct is measured in the same way in different target groups and is a necessary condition for the valid interpretation of interindividual differences (Vandenberg and Lance 2000). Only two community studies have investigated measurement invariance for the ICU, across gender (Essau et al. 2006) or gender and age (Ciucci et al. 2014). This is a critical gap in the literature considering that many clinicians may use the ICU in more severe justice-involved youth, which our study set out to fill.

Criterion Validity

More than 125 studies have demonstrated positive associations between psychopathic or CU traits and antisocial behavior in adolescents (see the comprehensive review in Frick et al. 2014), the vast majority using measures other than the ICU. Prior ICU validation studies (Roose et al. 2010; Ciucci et al. 2014; Fanti et al. 2009; Essau et al. 2006) and other studies using the ICU (e.g. Frick et al. 2003; Kimonis et al. 2007) found similar associations for ICU total, Callousness, and Uncaring subscale scores, mostly in the .35 to .45 range, and weaker ones for the Unemotional score, i.e. around .10 or below. To the exception of one study including the number of school formal warnings and peer-reported bullying (which observed correlations around .20; Ciucci et al. 2014), all studies measured antisocial behavior exclusively through self-report, so that the observed correlations may be inflated due to shared method variance. The present study aims at confirming the relation between ICU scores and antisocial behavior using a multi-method design combining self- and staff-reports.

CU traits are also moderately to highly associated with aggressive behavior (see the review by Muñoz and Frick 2012). Distinguishing between reactive aggression, i.e. "hot" aggression in retaliation to a perceived threat, and proactive aggression, i.e. "cold" goal-directed or instrumental aggression (Dodge and Coie 1987), studies using other instruments have shown that, among adolescents, psychopathic or CU traits are specifically associated with proactive (Kerig and Stellwagen 2010; Vitacco et al. 2006; Murrie et al. 2004) or with both types of aggression (Raine et al. 2006; Penney and Moretti 2010; Frick et al. 2003). Studies using the ICU found low to moderate associations with reactive aggression and stronger ones with proactive aggression for the total, Callousness and Uncaring scores, and null associations for the Unemotional score (Fanti et al. 2009; Marsee et al. 2014;

Stickle et al. 2012; Kimonis et al. 2008b). In summary, there is promising evidence of criterion validity regarding antisocial behavior and aggression for most ICU scores. An understudied issue is whether these associations are similar for boys and girls, for younger and older adolescents, and for community and institutionalized samples, and the present study aims to fill this gap in knowledge.

Deficits in responding to others' emotional distress cues theoretically fail to inhibit aggressive and antisocial actions against others, thus explaining high levels of these behaviors among youth with CU traits (Blair 1995). Adolescents with CU traits present with specific neurocognitive impairments in processing distress signals displayed through facial expressions and paraverbal cues (Blair et al. 2005; Stevens et al. 2001) or body language (Muñoz 2009). Most prior studies found impaired fear (and sometimes also sadness) recognition in community (Dadds et al. 2008; Dadds et al. 2006; Jones et al. 2009; A. A. Marsh et al. 2008; Blair and Coles 2000; Blair et al. 2001), high-risk (Stevens et al. 2001), or clinical (Fairchild et al. 2009) adolescents scoring high on CU traits. In contrast, their expression recognition was not impaired for emotions of joy, surprise, or anger. Most of these studies measured CU traits using the APSD or the YPI, and none of them used the ICU. The present study aims to test the criterion validity of the ICU in relation to a measure of emotional recognition.

Study Aims

In summary, the aims of the present study were to investigate 1) group differences in ICU scores across gender, age, and community versus institutionalized setting, as well as differences between the reference values of the present study and those of prior ICU validation studies using other LV/CC in similar samples; 2) the factor structure of the ICU and its stability across gender, age, and setting; 3) associations between ICU scores and relevant criterion measures (i.e. psychopathy, antisocial behavior, proactive aggression, misbehavior in residential settings, deficits in recognizing distress-related emotions) and their stability across gender, age, and setting.

Methods

Participants

Community Sample This sample comprised 397 adolescents (62 % girls) from the general population, with a mean age of 15.8 years (*S.D.* =1.9). Youths were recruited from seven public schools in the French-speaking part of Switzerland. Most of the participants were of Swiss nationality (83 %),

with French as their first language (85 %). Since all of the public schools provided students with an education free of charge, they did not discriminate on the basis of income, admitting adolescents from diverse socioeconomic backgrounds. However, there was a predominance of middle-class families with few of low SES (6 % low, 62 % middle, 32 % high SES; based on father's and mother's occupation).

Institutionalized Sample This sample comprised 164 institutionalized adolescents recruited as part of a larger study from youth welfare or juvenile justice institutions in the French-speaking part of Switzerland. Their mean age was 15.0 years (*S.D.* =2.0) and 70 % were boys. Further, 63 % were Swiss and 71 % had French as a first language. Half (51 %) had a current diagnosis of conduct disorder. As in the community sample, the majority came from middle-class families, however there were few high-SES families and one-fourth were of low SES (24 % low, 63 % middle, 13 % high SES; based on the same criteria as in the community sample).

Procedure

Community Sample The principals of the selected schools provided formal agreement to participate in the study. Students and their parent(s) participated after providing written informed consent. Pupils willing to participate (99.8 %) completed questionnaires individually and anonymously during a class period, within a single session lasting between 45 min and 2 h, under the supervision of research psychologists.

Institutionalized Sample The adolescent and his/her legal guardian (parents and/or judge) provided written informed consent. Adolescents willing to participate (97.7 %) completed questionnaires anonymously either 1) during an individual session with a research psychologist or a member of the institution staff, or 2) during a class period, under the supervision of research psychologists. Data collection, which included a number of other measures that were not used in the present study, lasted between 2 and 6 h.

Measures

We collected standard demographic information along with responses to three self-report questionnaires in the full sample, as well as additional measures described below in some subsamples.

Callous-Unemotional Traits The *Inventory of Callous-Unemotional Traits* (ICU; Frick 2004) has been extensively described in the introduction. No French version was available at the time of the study. Thus, we performed the translation following the cross-cultural research guidelines (Brislin 1970). The total score and the 3 scale scores were computed using the mean of the corresponding items.

The *Youth Psychopathic Traits Inventory* (YPI; Andershed et al. 2002) is a 50-item self-report questionnaire asking participants to estimate the degree to which each individual item applies to them, using a 4-point scale from 1=*does not apply at all* to 4=*applies very well.* It is composed of three scales: the Interpersonal scale measuring grandiosity and manipulative style, the Affective scale assessing restricted emotional experience and lack of remorse/empathy, and the Lifestyle scale measuring impulsivity and irresponsibility. This study focused on the Interpersonal and Affective scales, since they are conceptually similar to CU traits as assessed by the ICU. The YPI has shown good psychometric properties in prior research with adolescents (Skeem and Cauffman 2003). We used the French version (D'Acremont et al. 2002), which has also been found to be psychometrically sound (Pihet et al. 2014), with Cronbach alphas of .91 for the Interpersonal scale and .84 for the Affective scale in our sample.

Conduct Disorder Diagnosis In the institutionalized sample, the *Kiddie-SADS-Present and Lifetime Version* (K-SADS; Chambers et al. 1985) was used to ascertain the presence of conduct disorder (CD). This reliable semi-structured interview evaluates current and past episodes of psychopathology in children and adolescents (including CD), according to *DSM-IV* criteria (American Psychiatric Association, 1994). A trained research psychologist conducted the K-SADS during an individual interview with the adolescent.

Antisocial Behavior The *Youth Self-Report (YSR)* (Achenbach 1991) lists 113 specific problems commonly found in children and adolescents and has been validated in a sample of 11- to 18-year olds. It consists of two scales that reflect externalizing and internalizing problems. In the present study, we used only the *externalizing scale* (30 items), which includes aggressive and delinquent behavior subscales. Items are scored from 0=*not true* to 2=*often true*, judged over the last 12 months. YSR scores have demonstrated good reliability and validity (Vreugdenhil et al. 2006), acceptable test-retest reliability, and discriminant and convergent validity (Achenbach 1991). The French version of the YSR was used (Vermeersch and Fombonne 1997), with Cronbach alphas of .83 for aggressive and .81 for delinquent behavior in our sample.

Reactive and Proactive Aggression A subsample of the full sample described above (*n*=210),[1] including 118 community

---

[1] This subsample did not differ significantly from the remainder of the sample on ICU or YSR scores. However, they were significantly younger.

and 92 institutionalized adolescents (mean age 14.7 years, S.D. =1.1; 62 % boys), completed the *Reactive and Proactive Aggression Questionnaire* (RPQ; Raine et al. 2006). The RPQ is a 23-item questionnaire assessing reactive (11 items) and proactive (12 items) aggression in children and adolescents. The response scale ranges from 0=*never* to 2=*often*. The Cronbach alphas in our sample were .85 for reactive and .89 for proactive aggression.

Self- and Staff-Reported Antisocial Behavior in Daily Life A second subsample, composed exclusively of institutionalized adolescents ($n=88$[2]; mean age 14.8 years, S.D. =1.1; 76 % boys) housed in a residential facility (youth welfare or juvenile detention center), participated in a broader assessment in which staff and adolescents rated, four times per day across eight consecutive days for the previous four-hour time period, 1) the intensity of the participants' antisocial behaviors[3] on a scale from 0=*not at all* to 100=*extremely*, and 2) the number of antisocial behaviors presented out of list of 28. The average intensity and frequency of antisocial behaviors across the eight-day period was computed for adolescent and staff reports.

Recognition of Facial Expressions of Emotions A third subsample ($n=228$),[4] which included 120 community and 108 institutionalized adolescents (mean age 14.7 years, S.D. =1.0; 61 % boys), completed the Emotional Expression Multimorph Task (Blair et al. 2001). This computerized alternative forced choice task assesses recognition of six basic emotions (joy, surprise, anger, fear, sadness and disgust) using images taken from the empirically validated pictures of facial affect series (Ekman and Friesen 1976). Each of the 36 stimuli is presented as a continuum composed of 40 morphed images, viewed on a computer screen as a neutral face gradually morphing into one of the six prototypical expressions. Participants had to click on the emotion name printed at the bottom as soon as they recognized the expression.

Statistical Analyses

To test our first aim of examining *group differences*, we used ANOVA for the ICU total score across gender, age and setting (community versus institutionalized), and repeated measures ANOVA for the three subscale scores with subscale as the within-factor variable. To maximize the comparability of our

results with those of previous research (Essau et al. 2006), we examined differences between the same three age groups: 13–14 years, 15–16 years, and 17–18 years. Due to the small number of girls ($n=5$) in the 17–18 year old institutionalized group, we first included age and gender as between factors for analyses in the community sample. Then, only significant effects were included in analyses conducted in the full sample, with group (community versus institutionalized) as a between variables factor, and in additional analyses in the institutionalized group, exploring the effect of CD (with versus without).

To test our second aims of examining the *factor structure* of the ICU and its stability across groups, we first used CFAs with maximum likelihood estimation (in AMOS 20) to compare three a priori models (one-factor, three correlated factors, and a three-factor bifactor model) in the complete sample. The relative fit of the three models was evaluated using the following indices (for comparison purposes, we retained the indices provided in the previous studies): 1) the chi-square/degrees of freedom ratio (Raykov 1998) for which values lower than 5 and 3 respectively indicate acceptable and good fit (Milfont and Fischer 2010), 2) the Root Mean Square Error of Approximation, RMSEA (Hu and Bentler 1998), for which values lower than .08 and .05 represent acceptable and good fit, respectively (Schermelleh-Engel and Moosbrugger 2003), 3) the Bentler Comparative Fit Index (CFI), for which values higher than .80 and .90 represent acceptable and good fit, respectively (Hu and Bentler 1999), and 4) the Akaike Information Criterion (AIC), for which lower values indicate a better fit.

We then tested the *measurement invariance*, which essentially investigates whether the same constructs are measured in the same way in the compared groups and is a necessary prerequisite for meaningful group comparisons; to test this core psychometric property, we used the multigroup CFA, which allows testing different levels of invariance (e.g. Vandenberg and Lance 2000; Milfont and Fischer 2010). The lowest level is *configural invariance*, which postulates that the factor structure is the same across groups and is verified when the same model provides the best fit of the data in the compared subsamples. It was tested by comparing the fit of the three models, estimated separately in the community and institutionalized samples, for boys and girls, as well as for adolescents aged less than 16 and 16 years and over. The next level is *metric (or measurement weights) invariance*, which postulates that the factor loading are equal across groups, in other terms that the constructs are manifested in the same way across groups. The last level is *scalar (or measurement intercepts) invariance*, which postulates that the latent intercepts are equal across groups, in other terms that

---

[2] As expected, this institutionalized subsample scored higher than the rest of the sample on all YSR and ICU scores except for the Unemotional scale. They were also significantly younger.

[3] Two items assessed aggressive behavior (physical and verbal) and one item assessed rule-breaking behavior.

[4] As expected, this subsample, which included nearly half of institutionalized adolescents, scored significantly higher than the rest of the sample on YSR scores as well as on the ICU Total score and Unemotional and Callousness subscale scores; they were also significantly younger.

all observed variables indicate the same differences between the groups.[5] Metric and scalar invariance are confirmed by 1) an acceptable fit of the model with the corresponding parameters constrained to be equal in the compared groups, and 2) a difference in the Tucker-Lewis Index (ΔTLI) less than or equal to 0.05 between the unconstrained and constrained models (Little 1997).

Regarding the required sample size in CFA (for a critical review, see H. W. Marsh et al. 1998, pp.182–183), the most general recommendation is to have at least $N \geq 100$ and preferably $N \geq 200$ (e.g. Vandenberg and Lance 2000), and specific recommendations include having 10 times as many participants as observed variables (e.g. Meehan and Stuart 2007), which was the case for all analyses in the present study except for the institutionalized group ($n=164$). Internal consistency was assessed using Cronbach's alpha.
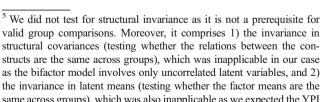
To test our third aim of examining associations between ICU scores and relevant criterion measures, we used Pearson correlations. Fisher's Z test was used to test for significant differences between correlations for boys versus girls, adolescents under versus over 16 years, and community versus institutionalized adolescents. Since multiple pairs of correlations were compared ($n=48$), we applied a Bonferroni correction (α was set at .001). As SES showed no significant association with the ICU scores, it was not controlled for in the analyses.

Results

Aim 1: Differences in ICU Scores Across Groups

As shown in Table 1, for all ICU scores, age had no significant main effect, while boys scored somewhat higher than girls ($F_{(1,543)}=11.99$, $p=.001$, $\eta^2=.02$), and institutionalized adolescents scored substantially higher than community adolescents ($F_{(1,543)}=38.02$, $p<.001$, $\eta^2=.07$); among institutionalized adolescents, those with CD scored significantly higher on CU traits than those without CD ($F_{(1,154)}=21.56$, $p<.001$, $\eta^2=.12$). All adolescents had higher scores on the Unemotional scale, intermediate on Uncaring, and lower ones on the Callousness scale ($F_{(2,381)}=717.35$, $p<.001$, $\eta^2=.79$). No interaction effect was significant. Reference values for all groups are provided in Table 1.

---

[5] We did not test for structural invariance as it is not a prerequisite for valid group comparisons. Moreover, it comprises 1) the invariance in structural covariances (testing whether the relations between the constructs are the same across groups), which was inapplicable in our case as the bifactor model involves only uncorrelated latent variables, and 2) the invariance in latent means (testing whether the factor means are the same across groups), which was also inapplicable as we expected the YPI to discriminate community and institutionalized adolescents, as well as between boys and girls.

We further investigated whether specific reference values are needed for different LV/CC. In our sample, using the French version of the ICU, Swiss and foreign participants did not differ significantly on the ICU scores (total score: $F_{(1,493)}=0.77$, $p=.382$; three subscale scores: $F_{(1,493)}=0.62$, $p=.431$). We also conducted one-sample t-tests comparing the scores obtained in our sample or subsamples to the means of groups similar in gender, age, and setting assessed with other LV/CC. After Bonferroni correction, 12 out of the 24 compared means differed significantly (see Appendix 1), confirming the need for reference values that are specific to the LV/CC.

Aim 2: ICU Factor Structure and its Stability Across Samples

The results of the CFA conducted in the full sample ($N=561$) indicated that the three-factor bifactor model best fit the data according to all fit indices (see Table 2, first block), with no differences between the 22- and 24-item versions. The absolute fit of the bifactor model was good according to χ^2/df and RMSEA, and acceptable according to CFI.

CFAs conducted separately for the community and institutionalized samples, for boys and girls, and for adolescents under and over 16, confirmed the superior fit of the three factor bifactor model in all samples (see Table 2), with acceptable to good absolute fit indices, supporting configural invariance. Metric and scalar invariance were also confirmed across gender, age, and setting, as indicated by the excellent fit of the multigroup model according to χ^2/df and RMSEA, and acceptable fit according to CFI (see Table 2), as well as the ΔTLI between the unconstrained and constrained model below 0.05 (for metric invariance: 0.017, 0.007 and 0.006, respectively; for scalar invariance: 0.018, 0.003 and 0.001, respectively).

Regarding internal consistency, in the full sample as well as in most subsamples, Cronbach's alphas were good for the total score (.79 in the full sample, and .76 to .80 in subsamples), satisfactory for the Callous (.72 in the full sample, and .66 to .78 in subsamples) and Uncaring (.73 in the full sample, and .69 to .76 in subsamples) subscales, and acceptable for the Unemotional subscale (.65 in the full sample, and .61 to .68 in subsamples), which could be at least in part due to the small number of items on this subscale. These values were highly similar to those reported in previous validation studies in adolescents (.77 to .83 for the total score; .66 to .80 for Callousness; .72 to .81 for Uncaring; .53 to .73 for Unemotional). Only a recent validation study conducted in a predominantly female college sample (Kimonis et al. 2012) found higher alphas for the Unemotional scale, and lower values for the Callous scale, which might be due to floor effects on the items of this scale.

Table 1 Mean, (standard deviation) and effect sizes (d) for the total and subscale scores (average across items) of the Inventory of Callous-Unemotional Traits, in community and institutionalized boys and girls, with and without a conduct-disorder (CD) diagnosis

	Community (n = 389)		Institutionalized (n=158)		Institutionalized-No CD[a] (n=79)		Institutionalized-With CD[b] (n=79)		Boys/Girls	Community/ Institutionalized	Institutionalized No CD/CD
	Boys (147)	Girls (242)	Boys (111)	Girls (47)	Boys (55)	Girls (24)	Boys (56)	Girls (23)	d	d	d
Total score	0.98 (0.3)	0.82 (0.3)	1.14 (0.4)	1.07 (0.4)	1.02 (0.3)	0.90 (0.3)	1.25 (0.4)	1.24 (0.5)	0.52	0.69	0.76
Unemotional	1.51 (0.5)	1.41 (0.6)	1.65 (0.7)	1.55 (0.7)	1.68 (0.5)	1.42 (0.6)	1.62 (0.7)	1.69 (0.6)	0.67	0.77	0.88
Uncaring	1.21 (0.4)	1.07 (0.5)	1.36 (0.6)	1.30 (0.6)	1.18 (0.6)	1.19 (0.6)	1.53 (0.5)	1.42 (0.7)	0.27	0.40	0.54
Callous	0.57 (0.3)	0.36 (0.3)	0.74 (0.4)	0.68 (0.5)	0.60 (0.3)	0.46 (0.3)	0.89 (0.4)	0.90 (0.6)	0.18	0.28	0.06

Aim 3: Associations Between ICU Scores and Criterion Measures

Psychopathic Traits and Externalizing Behavior As expected, correlations between the ICU total score and the Affective and Interpersonal scales of the YPI were large (see Table 3). Correlations were substantial for the Callousness and Uncaring scales, and modest for the Unemotional scale. This pattern of correlations was remarkably stable across subsamples; all differences between subsamples were not significant according to Fisher's Z test.

In line with our hypotheses, the ICU total score also correlated strongly with self-reported delinquent behavior, and moderately with aggressive behavior. Again, correlations were substantial for the Callousness and Uncaring scales, and modest for the Unemotional scale, with a highly stable pattern across subsamples (all Fisher's Z tests were not significant).

Reactive and Proactive Aggression In line with Fanti et al. (2009), hierarchical linear regression analyses were conducted, with RPQ Proactive and Reactive aggression scores as outcomes, controlling for gender in the first step and for externalizing problems (YSR score) in the second step so as to avoid spurious effects due to the well-known association between CU traits and externalizing problems. We also controlled for the other form of aggression (e.g. reactive aggression when predicting proactive aggression) in the third step so as to isolate each specific form of aggression from the general level of aggression. In one model, the total ICU score was entered in the fourth step, and in the second model, the three subscale scores were entered.

As presented in Table 4, ICU total, Callousness and Uncaring scores contributed significantly to the prediction of proactive aggression, over and above gender, externalizing problems, and general aggression, while none of the ICU scores predicted reactive aggression. However, the Unemotional subscale failed to predict proactive aggression.

Self- and Staff-Reported Antisocial Behavior in Daily Life Given that most distributions were markedly asymmetric, we computed Spearman's rank correlations between each of these variables and the ICU scores, applying a Bonferroni correction. As presented in Table 5, adolescents with higher Callousness and Uncaring scores reported substantially more intense and more frequent antisocial behavior during their stay in a residential facility. Staff members also reported significantly more intense and non-significantly more frequent antisocial behavior among those with higher Uncaring scores. Again, the Unemotional scores were not significantly associated with this measure of antisocial behavior. Highly similar

Table 2 Fit indices for confirmatory factor analyses comparing three models of the structure of the Inventory of Callous-Unemotional Traits (ICU)

Model	χ²/df[a]		CFI[b]		RMSEA[c]		AIC[d]	
Present study—All	24i	22i	24i	22i	24i	22i	24i	22i
1 factor	6.4	6.9	.51	.53	.098	.103	1755	1578
3 correlated factors	4.3	4.6	.70	.72	.077	.080	1219	1083
3-factor bifactor	**3.1**	**3.1**	**.83**	**.85**	**.061**	**.062**	**898**	**761**
Kimonis (2008) bifactor	**2.2**	**1.8**	**.79**	**.87**	**.070**	**.060**		
Essau (2006) bifactor-24i[e]	**8.0**		**Not indicated**		**.100**			
Fanti (2009) bifactor -24i[e]	**2.2**		**.82**		**.058**			
Roose (2009) bifactor-24i[e]	**3.0**		**.92**		**.070**			
Feilhauer (2012) bifa-24i[e]	**3.1**		**.82**		**.077**			
Ciucci (2013) bifactor-24i[e]	**2.2**		**.87**		**.050**			
Community/Institutional.	Com[f]	Inst[g]	Com[f]	Inst[g]	Com[f]	Inst[g]	Com[f]	Inst[g]
1 factor	4.8	2.9	.49	.36	.098	.107	1351	870
3 correlated factors	3.5	2.2	.67	.61	.079	.085	1012	693
3-factor bifactor	**1.8**	**2.6**	**.81**	**.76**	**.063**	**.069**	**772**	**599**
metric invariance	2.1		.82		.044			
scalar invariance	2.1		.80		.044			
Boys/Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1 factor	3.6	4.0	.46	.50	.101	.102	1055	1152
3 correlated factors	2.8	2.8	.64	.70	.083	.079	839	852
3-factor bifactor	**2.2**	**2.4**	**.78**	**.79**	**.068**	**.069**	**691**	**738**
metric invariance	2.2		.78		.047			
scalar invariance	2.3		.75		.049			
Under/Over 16	<16	≥16	<16	≥16	<16	≥16	<16	≥16
1 factor	4.2	3.3	.52	.51	.100	.102	1201	985
3 correlated factors	3.2	2.4	.67	.72	.083	.078	950	740
3-factor bifactor	**2.4**	**2.0**	**.80**	**.81**	**.067**	**.066**	**747**	**647**
metric invariance	2.2		.80		.046			
scalar invariance	2.2		.79		.046			

Based on the 24 or 22 items of the ICU, for the complete sample (All; N=561), as well as for the community (n=397) and institutionalized (n=164) samples, for boys (n=259) and girls (n=291), and adolescents under (n=322) and over 16 (n=228), as well as in previous studies (only the first author is listed); [a] Chi-square to degrees of freedom ratio; [b] Comparative Fit Index; [c] Root Mean Square Error of Approximation; [d] Akaike Information Criterion; [e] 24 items; [f] Community; [g] Institutionalized

patterns of correlations were observed for adolescents with and without CD.

Recognition of Facial Expressions of Emotions Joy was excluded from analyses due to nearly perfect accuracy for all participants. Since the accuracies for the five other emotions were modestly correlated (average $r=.12$, ranging from -.06 to .32), we conducted analyses on the average accuracy as well as on each emotion separately. In addition, as girls were globally more accurate than boys ($F_{(5,230)}=2.86$, $p=.016$, $\eta^2=.059$, average accuracy $.72\pm.10$ and $.69\pm.11$, respectively), in particular regarding anger ($F_{(1,234)}=9.15$, $p=.003$, $\eta^2=.038$, $.70\pm.17$ and $.63\pm.18$, respectively) and sadness ($F_{(1,234)}=4.49$, $p=.035$, $\eta^2=.019$, $.75\pm.20$ and $.69\pm.22$, respectively), gender was controlled in all analyses. All correlations between ICU scores and accuracy were non-significant and smaller than .10, with the exception of those between the Uncaring subscale and the average accuracy ($r=-.15$, $p=.026$), as well as the accuracy for sadness ($r=-.14$, $p=.043$) and fear ($r=-.18$, $p=.009$). This pattern of results was observed in both the CD and non-CD groups.

In the interest of examining within a single model the respective contributions of CD and CU traits to accuracy, and to investigate possible interaction effects between them, we conducted a MANOVA with accuracy as the dependant variable, CD (present versus absent) and CU traits (high versus low, based on median split) as between factor variables, and type of emotion (surprise, anger, fear, sadness, and disgust) as a within-factor variable. While CD had no significant main effect, adolescents higher on CU traits were found to be significantly less accurate than their low-CU counterparts (main effect of CU traits: $F_{(1,223)}=4.73$, $p=.031$, $\eta^2=.021$), with an interaction effect between CD and CU traits ($F_{(1,223)}=4.79$, $p=.030$, $\eta^2=.021$): among non-CD adolescents, those with low and high CU traits had similar average accuracy levels ($.71\pm.09$ and $.70\pm.10$, respectively), while among CD adolescents, those with lower CU traits showed a substantially higher average accuracy ($.76\pm.09$) than their high-CU counterparts ($.68\pm.14$). In addition, the interaction between emotion and CU traits was significant ($F_{(3,221)}=3.80$, $p=.011$, $\eta^2=.049$), pointing to a lower accuracy in the high-CU group specific to distress-related emotions (fear, sadness and disgust), as illustrated in Fig. 1. This was confirmed by the significant contrast comparing the mean accuracy across fear, sadness and disgust as opposed to anger and surprise ($F_{(1,223)}=8.72$, $p=.003$, $\eta^2=.038$). The triple interaction was not significant. In summary, controlling for gender, adolescents with higher CU traits were on average significantly less accurate at recognizing distress-related emotions than their low-CU peers, and their lower accuracy was more pronounced when they additionally met criteria for a CD diagnosis.

Discussion

The main aim of this study was to examine the psychometric properties of the ICU in a French-speaking Swiss sample and to examine measurement invariance across various groups of adolescents. Our study contributes three novel findings. First, our results confirmed the stability of ICU scores across age, gender, and community/institutionalized groups for the measure's factor structure and associations with several criterion measures. The ICU can thus be considered to measure the same constructs in the same way across all of the groups studied. Second, it offers the first evidence that the ICU discriminates between community and institutionalized adolescents with and without CD in a sample including boys and girls. Third, our results lend additional support to the construct validity of the ICU, as most ICU scores presented the expected relations with other measures in the same nomological net, to the exception of the Unemotional subscale. Most ICU scores statistically predicted 1) scores on affective and interpersonal subscales of a measure of juvenile psychopathy, 2) proactive aggression, over and above gender and antisocial behavior, 3) self- and staff-rated intensity of misconduct during residential stay, and 4) reduced accuracy in recognizing distress-related emotions. However, the Unemotional scale score failed to correlate with most criterion measures.

Our findings extend previous evidence obtained in either community (Essau et al. 2006; Fanti et al. 2009; Roose et al. 2010) or institutionalized samples (Kimonis et al. 2008b), and recently in a male-only mixed sample (Feilhauer et al. 2012), by confirming that the three-factor bifactor model best fits the data in all adolescent groups. Compared to previous studies (see Table 3), fit was largely better than that found by Essau et al. (2006), and comparable to the best fit observed by Kimonis et al. (2008b) for the 22-item version, or by Roose et al. (2010), Fanti et al. (2009) or Feilhauer et al. (2012) for the full 24-item version. In addition, our results support the measurement invariance of ICU scores across gender (as also found by Essau et al. 2006), as well as across age and community/institutionalized groups. This novel evidence that the factor structure, as well as measurement weights and intercepts, were comparable between these different samples, suggests that the ICU is able to capture the same psychological construct in the same way across various groups of adolescents. The present study provides the first comprehensive support for this necessary condition regarding the valid interpretation of interindividual differences. In addition, it should be noted that the currently available evidence regarding the factor structure of the ICU is remarkably convergent, and encourages clinicians and researchers to rely on the ICU total score and at least two of three subscale scores, regardless of the language version (English, Dutch, German, Greek or French) and adolescent target group (i.e., boys or girls, younger or older, and community or institutionalized adolescents).

Table 3 Pearson correlation coefficients between the Inventory of Callous-Unemotional Traits (ICU) total score and three subscale scores and 1) the Youth Psychopathic Traits Inventory (YPI) Affective and Interpersonal scale scores, as well as 2) the Child Behavior Checklist (CBCL) Delinquent and Aggressive behavior scores, for the complete sample and the subsamples

		All (548)	Boys (254)	Girls (289)	<16 (320)	≥16 (226)	Com[a](394)	Inst[b] (156)
YPI-Affective	ICU-TS[c]	**0.62*****	0.62***	0.54***	0.58***	0.67***	0.61***	0.51***
	ICU-Une[d]	**0.29*****	0.24***	0.28***	0.19**	0.41***	0.30***	0.18*
	ICU-Unc[e]	**0.45*****	0.49***	0.39***	0.48***	0.42***	0.44***	0.40***
	ICU-Cal[f]	**0.58*****	0.55***	0.50***	0.55***	0.62***	0.57***	0.45***
YPI-Interpersonal	ICU-TS[c]	**0.47*****	0.45***	0.41***	0.45***	0.49***	0.40***	0.42***
	ICU-Une[d]	**0.14*****	0.18**	0.06	0.09	0.20**	0.09	0.13
	ICU-Unc[e]	**0.40*****	0.38***	0.38***	0.41***	0.40***	0.32***	0.43***
	ICU-Cal[f]	**0.44*****	0.35***	0.43***	0.43***	0.44***	0.43***	0.28***
CBCL-Delinquent	ICU-TS[c]	**0.51*****	0.48***	0.51***	0.49***	0.54***	0.46***	0.44***
	ICU-Une[d]	**0.16*****	0.14	0.17***	0.09	0.26***	0.14**	0.10
	ICU-Unc[e]	**0.40*****	0.38***	0.40***	0.43***	0.37***	0.34***	0.41***
	ICU-Cal[f]	**0.50*****	0.45***	0.53***	0.50***	0.51***	0.48***	0.35***
CBCL-Aggressive	ICU-TS[c]	**0.35*****	0.36***	0.37***	0.40***	0.29***	0.25***	0.41***
	ICU-Une[d]	**−0.01**	−0.01	−0.01	−0.02	0.00	−0.05	−0.02
	ICU-Unc[e]	**0.38*****	0.39***	0.41***	0.41***	0.38***	0.28***	0.48***
	ICU-Cal[f]	**0.32*****	0.32***	0.36***	0.39***	0.22***	0.25***	0.30***

[a] Community; [b] Institutionalized; [c] ICU Total score; [d] ICU Unemotional; [e] ICU Uncaring; [f] ICU Callous

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

Table 4 Hierarchical linear regression analyses predicting proactive and reactive aggression from the ICU total and subscale scores, after controlling for gender, externalizing problems and the other aggression form (n=210)

	Proactive Aggression			Reactive Aggression		
	B	β	ΔR^2	B	β	ΔR^2
Step1: Gender	-.17	-.26***	.07***	-.05	-.06	.00
Step2: CBCL-Externalizing	.68	.62***	.39***	.83	.61***	.38***
Step3: Other aggression form	.29	.37***	.08***	.52	.09***	.10***
Step4a: ICU-Total score	.07	.10*	.01*	-.07	-.08	.01
Step4b: ICU-Subscale scores			.05***			.00
Callous	.18	.21***		-.05	-.05	
Uncaring	.07	.12*		.02	.03	
Unemotional	.00	.00		-.03	-.04	

Taken together, these findings provide strong support to the way the ICU would most likely be used in practice – i.e., drawing upon the total score as an indicator of the overall level of callous-unemotional traits, paired with an analysis of the subscale scores to inform different domains of a youth's functioning.

This study also offered the first confirmation that the ICU total and three subscale scores evidenced similar associations with criterion measures in all adolescent groups, lending further support to the criterion and construct validity of most ICU scores, and confirming the inconsistent findings regarding the Unemotional score. The criterion validity was confirmed based on associations with another commonly-used measure of juvenile psychopathy, the YPI, and with scores on measures of antisocial behavior (multi-method), proactive aggression, and emotion recognition. This last result additionally provides an important contribution to the current debate about the deficits in emotion recognition associated to psychopathy, illustrated by the contradictory conclusions of two recent meta-analyses on the topic: While the first concluded to an effect of small size as to fear and sadness (and possibly also disgust and happiness for studies using nonverbal answer modalities as in our case; Wilson et al. 2011), the second concluded to a pervasive emotion recognition deficit (though observing a significant association with CU traits only for the mean of all emotions and fear; Dawel et al. 2012). Interestingly, these studies did not address a possible interaction between CD and CU traits, which our findings suggest to be highly relevant. Indeed, we observed the emotion recognition deficits associated with CU traits only in adolescents with CD, confirming previous evidence regarding deficits in the emotional processing of distress cues in community boys and girls (Kimonis et al. 2006) as well as in detained adolescent boys (Kimonis et al. 2008a). These findings point to the importance of taking into account externalizing behavior, as it may be a key moderator of the relationship between CU traits and emotional deficits.

Notably, in line with previous findings, the *Unemotional* scale consistently demonstrated low to null correlations with criterion measures, as also recently documented in preschoolers based on parent and teacher ratings of the ICU (Ezpeleta et al. 2013). This effect might be attributed to the lower internal consistency of this scale, which we also observed (α=.65), possibly resulting from its small number of items. Alternatively, these low correlations could be due to the focus of the Unemotional scale on restricted emotional expression, using general assertions such as "I hide my feelings to others". This focus may fail to capture essential features of the emotional life that characterizes adolescent psychopathy. For example, in this regard, the YPI assesses 1) emotional impassibility, focusing on reduced fear and sadness experience, and 2) emotional insensitivity, focusing on reduced emotional contagion and the perception of emotions as weaknesses. This limitation has also been noted by the experts in charge of the revisions to the CD diagnosis for DSM-5 (Frick and Moffitt 2010). These authors underline that an important feature of CU traits is the display of emotions in order to obtain things from others or manipulate them, and propose to explicitly include this characteristic in one of the three criteria for a diagnosis of CU traits, consequently renamed "shallow or deficient affect" instead of "unemotional" (ibid, p.3). However, it is also possible that questions related to emotional life are difficult to answer for adolescents growing up in families with limited emotional validation and little communication about feelings. In conclusion, taken together with previous evidence, our results suggest that the validity of the ICU might be further improved by either removing its Unemotional scale, or revising it, either through the addition of new items or through the replacement of its existing items with items that better capture the essence of shallow and deficient affect.

Regarding group differences, our results were consistent with the limited available evidence suggesting that boys score higher on the ICU total and subscale scores than girls (Essau

Table 5 Spearman's rank correlations between 1) the ICU total and subscale scores, and 2) the intensity and frequency of self- and staff-reported average daily antisocial behaviors (AB) across 8 days of Ecological Momentary Assessment (n=88; based on 2,313 self-report and 814 staff-report occasions)

	Total	Callous	Uncaring	Unemo.
Intensity of self-reported AB	.03	.36*	.44*	.04
Frequency of self-reported AB	.04	.45*	.43*	-.04
Intensity of staff-reported AB	.04	.14	.35*	.04
Frequency of staff-reported AB	-.03	.12	.23	.01

* *p*<.003 (Bonferroni corrected)

et al. 2006; Fanti et al. 2009). However, we observed no age differences in ICU scores, contrary to the only prior study that addressed this question (Essau et al. 2006). Notably, our results are in line with recently published evidence for substantially higher ICU scores in institutionalized male adolescents compared to their community peers (Feilhauer et al. 2012), although we found additional significant differences between institutionalized adolescents with and without a CD diagnosis, while these authors found none between their clinical and non-clinical offenders. These divergent findings might be explained by the more stringent criteria used in the present study (i.e., a formal diagnosis, compared to recruitment from a forensic facility). In summary, our findings regarding group differences corroborate the discriminant validity of the ICU, and suggest that further research in the same line, with other language versions and/or in other cultural contexts (LV/CC), is urgently needed. Indeed, comparisons of our group means with those reported in prior research using

Fig. 1 Mean accuracy for the recognition of five basic emotions in adolescents with low or high callous-unemotional (CU) traits, assessed with the Emotional Expression Multimorph Task (n=228)

different LV/CC pointed to substantial differences regarding ICU scores, as is also the case for other instruments such as the widely used Child Behavior Checklist (Rescorla et al. 2007). Thus, before the ICU can be used for screening purposes, additional evidence is definitely needed regarding gender, age and community/ institutionalized sample differences for different LV/CC.

These results should however be considered in light of some study limitations. Firstly, our demonstration that the criterion and construct validity of the ICU was stable across various adolescent groups has remained limited to concurrent self-report measures. More far-reaching conclusions would require its extension to predictive validity, as well as to additional measures unaffected by shared method variance, such as teacher/staff reports. Secondly, due to sample size constraints, we could not test for measurement invariance between institutionalized adolescents with and without a CD diagnosis, or in more specific groups such as girls with CD (Berkout et al. 2011), nor test for gender x age interactions in the institutionalized sample. In addition, due to practical constraints, some criterion measures were only assessed in a partial sample, which in some occasions (i.e. daily life antisocial behavior and recognition of facial expressions of emotions) scored relatively high on antisocial behavior and CU traits, compared to the rest of the sample. Further studies should thus confirm the generalizability of the reported findings to samples scoring lower on antisocial behavior and CU traits.

The present study, nonetheless, had a number of strengths. In particular, our sample covered a broad range of youth from well-functioning community adolescents to incarcerated juvenile offenders, and ranged in age from 13 to 18 years. It also comprised enough girls in the institutionalized sample to warrant meaningful group comparisons. In addition, the criteria used to assess the validity of ICU scores included multiple methods, such as ecological momentary assessment and a computerized task of facial emotion recognition, in addition to self-report. Our findings suggest that the ICU holds promises as a brief screening instrument for identifying antisocial adolescents at high risk due to their co-occurring CU traits.

References

Achenbach, T. M. (1991). *Manual for the youth self-report and 1991 profiles*. Burlington: University of Vermont, Department of Psychiatry.

Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non referred youths: A new assessment tool. In E. Blaauw & L. Sheridan (Eds.), *Psychopaths: Current International Perspectives* (pp. 131–158). The Hague: Elsevier.

Berkout, O. V., Young, J. N., & Gross, A. M. (2011). Mean girls and bad boys: Recent research on gender differences in conduct disorder. *Aggression and Violent Behavior, 16*(6), 503–511. doi:10.1016/j.avb.2011.06.001.

Blair, R. J. R. (1995). A cognitive developmental-approach to morality - Investigating the psychopath. *Cognition, 57*(1), 1–29. doi:10.1016/0010-0277(95)00676-P.

Blair, R. J. R., & Coles, M. (2000). Expression recognition and behavioural problems in early adolescence. *Cognitive Development, 15*(4), 421–434. doi:10.1016/s0885-2014(01)00039-9.

Blair, R. J. R., Colledge, E., Murray, L., & Mitchell, D. G. V. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology, 29*(6), 491–498. doi:10.1023/A:1012225108281.

Blair, R. J. R., Budhani, S., Colledge, E., & Scott, S. (2005). Deafness to fear in boys with psychopathic tendencies. *Journal of Child Psychology and Psychiatry, 46*(3), 327–336. doi:10.1111/j.1469-7610.2004.00356.x.

Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2006). Continuity and change in psychopathic traits as measured via normal-range personality: A longitudinal-biometric study. *Journal of Abnormal Psychology, 115*(1), 85–95. doi:10.1037/0021-843X.115.1.85.

Brislin, R. W. (1970). Back-Translation for Cross-Cultural Research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. doi:10.1177/135910457000100301.

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79*(5), 1185–1229. doi:10.1111/j.1467-8624.2008.01184.x.

Chambers, W. J., Puig-Antich, J., Hirsh, M., Paez, P., Ambrosini, P. J., Tabrizi, M. A., et al. (1985). The assessment of affective disorders in children and adolescents by semistructured interview. Test-retest reliability of the schedule for affective disorders and schizophrenia for school-age children, present episode version. *Archives of General Psychiatry, 42*(7), 696–702.

Ciucci, E., Baroncelli, A., Franchi, M., Golmaryami, F. N., & Frick, P. J. (2014). The association between callous-unemotional traits and behavioral and academic adjustment in children: Further validation of the Inventory of Callous-Unemotional Traits. *Journal of Psychopathology and Behavioral Assessment*. doi:10.1007/s10862-013-9384-z.

D'Acremont, M., Van der Linden, M., Axelson, H., Flykt, A., & Vonèche, J. (2002). French version of the Youth Psychopathic Inventory. *Unpublished manuscript.*

Dadds, M. R., Whiting, C., & Hawes, D. J. (2006). Associations among cruelty to animals, family conflict, and psychopathic traits in childhood. *Journal of Interpersonal Violence, 21*(3), 411–429.

Dadds, M. R., El Masry, Y., Wimalaweera, S., & Guastella, A. J. (2008). Reduced eye gaze explains "Fear Blindness" in childhood psychopathic traits. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(4), 455–463. doi:10.1097/CHI.0b013e31816407f1.

Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews, 36*(10), 2288–2304. doi:10.1016/j.neubiorev.2012.08.006.

Decuyper, M., De Bolle, M., De Fruyt, F., & De Clercq, B. (2011). General and maladaptive personality dimensions and the assessment of callous-unemotional traits in adolescence. *Journal of Personality Disorders, 25*(5), 681–701.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in childrens peer groups. *Journal of Personality and Social Psychology, 53*(6), 1146–1158. doi:10.1037/0022-3514.53.6.1146.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1*(1), 56–75. doi:10.1007/Bf01115465.

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment, 13*(4), 454–469. doi:10.1177/1073191106287354.

Ezpeleta, L., de la Osa, N., Granero, R., Penelo, E., & Domenech, J. M. (2013). Inventory of callous-unemotional traits in a community sample of preschoolers. *Journal of Clinical Child and Adolescent Psychology, 42*(1), 91–105.

Fairchild, G., Van Goozen, S. H. M., Calder, A. J., Stollery, S. J., & Goodyer, I. M. (2009). Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *Journal of Child Psychology and Psychiatry, 50*(5), 627–636. doi:10.1111/j.1469-7610.2008.02020.x.

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.

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

Frick, P. J. (2004). The inventory of callous-unemotional traits. *Unpublished manuscript.*

Frick, P. J., & Dickens, C. (2006). Current perspectives on conduct disorder. *Current Psychiatry Reports, 8*(1), 59–72. doi:10.1007/s11920-006-0082-3.

Frick, P. J., & Moffitt, T. E. (2010). A proposal to the DSM-V childhood disorders and the ADHD and disruptive behavior disorders work groups to include a specifier to the diagnosis of conduct disorder based on the presence of callous-unemotional traits. *American Psychiatric Association.*

Frick, P. J., & White, S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry, 49*(4), 359–375. doi:10.1111/j.1469-7610.2007.01862.x.

Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report

of delinquency. *Journal of Abnormal Child Psychology, 31*(4), 457–470. doi:10.1023/A:1023899703866.

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.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparametrized model mis-specification. *Psychological Methods, 3*, 424–453. doi:10.1037/1082-989X.3.4.424.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis : Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. doi:10.1080/10705519909540118.

Jones, A. P., Laurens, K. R., Herba, C. M., Barker, G. J., & Viding, E. (2009). Amygdala hypoactivity to fearful faces in boys with conduct problems and callous-unemotional traits. *American Journal of Psychiatry, 166*(1), 95–102. doi:10.1176/appi.ajp.2008.07071050.

Kerig, P. K., & Stellwagen, K. K. (2010). Roles of callous-unemotional traits, narcissism, and machiavellianism in childhood aggression. *Journal of Psychopathology and Behavioral Assessment, 32*(3), 343–352. doi:10.1007/s10862-009-9168-7.

Kimonis, E. R., Frick, P. J., Fazekas, H., & Loney, B. R. (2006). Psychopathy, aggression, and the processing of emotional stimuli in non-referred girls and boys. *Behavioral Sciences and the Law, 24*(1), 21–37. doi:10.1002/bsl.668.

Kimonis, E. R., Frick, P. J., Muñoz, L. C., & Aucoin, K. J. (2007). Can a laboratory measure of emotional processing enhance the statistical prediction of aggression and delinquency in detained adolescents with callous-unemotional traits? *Journal of Abnormal Child Psychology, 35*(5), 773–785. doi:10.1007/s10802-007-9136-1.

Kimonis, E. R., Frick, P. J., Muñoz, L. C., & Aucoin, K. J. (2008a). Callous-unemotional traits and the emotional processing of distress cues in detained boys: Testing the moderating role of aggression, exposure to community violence, and histories of abuse. *Development and Psychopathology, 20*(2), 569–589. doi:10.1017/s095457940800028x.

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Muñoz, L. C., et al. (2008b). Assessing callous-unemotional 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.

Kimonis, E. R., Branch, J., Hagman, B., Graham, N., & Miller, C. (2012). The psychometric properties of the inventory of callous-unemotional traits in an undergraduate sample. *Psychological Assessment.* doi:10.1037/a0029024.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*(1), 53–76.

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.

Marsee, M. A., Lau, K. S. L., & Lapre, G. E. (2014). Parent and adolescent report of the forms and functions of aggression: Associations with delinquency, CU traits, and dysregulation. *Child and Youth Care Forum, 43*(1), 27–39.

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*(2), 181–220. doi:10.1207/s15327906mbr3302_1.

Marsh, A. A., Finger, E. C., Mitchell, D. G. V., Reid, M. E., Sims, C., Kosson, D. S., et al. (2008). Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional traits and disruptive behavior disorders. *American Journal of Psychiatry, 165*(6), 712–720. doi:10.1176/appi.ajp.2007.07071145.

Meehan, J. C., & Stuart, G. L. (2007). Using structural equation modeling with forensic samples. *Criminal Justice and Behavior, 34*(12), 1560–1587.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111–121.

Muñoz, L. C. (2009). Callous-unemotional traits are related to combined deficits in recognizing afraid faces and body poses. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(5), 554–562. doi:10.1097/CHI.0b013e31819c2419.

Muñoz, L. C., & Frick, P. J. (2007). The reliability, stability, and predictive utility of the self-report version of the antisocial process screening device. *Scandinavian Journal of Psychology, 48*(4), 299–312. doi:10.1111/j.1467-9450.2007.00560.x.

Muñoz, L. C., & Frick, P. J. (2012). Callous-unemotional traits and their implication for understanding and treating aggressive and violent youths. *Criminal Justice and Behavior, 39*(6), 794–813. doi:10.1177/0093854812437019.

Muñoz, L. C., Frick, P. J., Kimonis, E. R., & Aucoin, K. J. (2008). Types of aggression, responsiveness to provocation, and callous-unemotional traits in detained adolescents. *Journal of Abnormal Child Psychology, 36*(1), 15–28.

Muñoz, L. C., Pakalniskiene, V., & Frick, P. J. (2011). Parental monitoring and youth behavior problems: Moderation by callous-unemotional traits over time. *European Child and Adolescent Psychiatry, 20*(5), 261–269.

Murrie, D. C., Cornell, D. G., Kaplan, S., McConville, D., & Levy-Elkon, A. (2004). Psychopathy scores and violence among juvenile offenders: A multi-measure study. *Behavioral Sciences and the Law, 22*(1), 49–67. doi:10.1002/bsl.573.

Penney, S. R., & Moretti, M. M. (2010). The roles of affect dysregulation and deficient affect in youth violence. *Criminal Justice and Behavior, 37*(6), 709–731. doi:10.1177/0093854810365446.

Pihet, S., Suter, M., Meylan, N., & Schmid, M. (2014). Factor structure of the Youth Psychopathic Traits Inventory in community and institutionalized adolescents: Using the total score, three scale scores, and/or 10 subscale scores. *Criminal Justice and Behavior, 41*(10), 1214-1231. doi:10.1177/0093854814540287.

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., et al. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*(2), 159–171.

Raykov, T. (1998). On the use of confirmatory factor analysis in personality research. *Personality and Individual Differences, 8*, 23–74. doi:10.1016/S0191-8869(97)00159-1.

Rescorla, L., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., et al. (2007). Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology, 75*(2), 351–358. doi:10.1037/0022-006X.75.2.351.

Roose, A., Bijtterbier, P., Decoene, S., Claes, L., & Frick, P. J. (2010). Assessing the affective features of psychopathy in adolescence: A further validation of the Inventory of Callous and Unemotional Traits. *Assessment, 17*, 44–57. doi:10.1177/1073191109344153.

Schermelleh-Engel, K., & Moosbrugger, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23–74.

Skeem, J. L., & Cauffman, E. (2003). Views of the downward extension: Comparing the youth version of the psychopathy checklist with the youth psychopathic traits inventory. *Behavioral Sciences and the Law, 21*(6), 737–770. doi:10.1002/bsl.563.

Stevens, D., Charman, T., & Blair, R. J. R. (2001). Recognition of emotion in facial expressions and vocal tones in children with psychopathic tendencies. *Journal of Genetic Psychology, 162*(2), 201–211.

Stickle, T. R., Marini, V. A., & Thomas, J. N. (2012). Gender differences in psychopathic traits, types, and correlates of aggression among adjudicated youth. *Journal of Abnormal Child Psychology, 40*(4), 513–525.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–69. doi:10.1177/109442810031002.

Vermeersch, S., & Fombonne, E. (1997). Le child behavior checklist: Résultats préliminaires de la standardisation de la version française. *Neuropsychiatrie de l'Enfance et de l'Adolescence, 45*, 615–620.

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, and Allied Disciplines, 50*(4), 471–481.

Vitacco, M. J., Neumann, C. S., Caldwell, M. F., Leistico, A. M., & Van Rybroek, G. J. (2006). Testing factor models of the psychopathy checklist: Youth version and their association with instrumental aggression. *Journal of Personality Assessment, 87*(1), 74–83.

Vreugdenhil, C., Van den Brink, W., Ferdinand, R., Wouters, L., & Doreleijers, T. (2006). The ability of YSR scales to predict DSM/DISC-C psychiatric disorders among incarcerated male adolescents. *European Child and Adolescent Psychiatry, 15*(2), 88–96.

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 and the Law, 27*(6), 910–928.

Wilson, K., Juodis, M., & Porter, S. (2011). Fear and loathing in psychopaths : A meta-analytic investigation of the facial affect recognition deficit. *Criminal Justice and Behavior, 38*(7), 659–668. doi:10.1177/0093854811404120.