International Journal of Law and Psychiatry 31 (2008) 241–252

Contents lists available at ScienceDirect

International Journal of Law and Psychiatry

Assessing callous–unemotional traits in adolescent offenders: Validation of the Inventory of Callous–Unemotional Traits

Eva R. Kimonis [a,*], Paul J. Frick [b], Jennifer L. Skeem [c], Monica A. Marsee [b], Keith Cruise [d], Luna C. Munoz [e], Katherine J. Aucoin [f], Amanda S. Morris [g]

[a] University of South Florida, United States
[b] University of New Orleans, United States
[c] University of California, Irvine, United States
[d] Fordham University, United States
[e] University of Central Lancashire, United Kingdom
[f] Jefferson Parish Human Services Authority, United States
[g] Oklahoma State University, United States

ARTICLE INFO

Keywords:
Assessment
Callous–unemotional traits
Aggression
Delinquency
Adolescents

ABSTRACT

The presence of callous–unemotional (CU) traits designates an important subgroup of antisocial youth. To improve upon existing measures, the Inventory of Callous–Unemotional Traits (ICU) was developed to provide an efficient, reliable, and valid assessment of CU traits in samples of youth. The current study tests the factor structure and correlates of the ICU scale in a sample (n=248) of juvenile offenders (188 boys, 60 girls) between the ages of 12 and 20 (M=15.47, SD=1.37). Confirmatory factor analyses are consistent with the presence of three independent factors (i.e., Uncaring, Callousness, and Unemotional) that relate to a higher-order callous–unemotional dimension. Also, CU traits overall showed associations with aggression, delinquency, and both psychophysiological and self-report indices of emotional reactivity. There were some important differences across the three facets of the ICU in their associations with these key external criteria.

1. Introduction

Callous–unemotional (CU) traits (e.g., lack of empathy, lack of guilt, poverty in emotional expression) are relatively stable across childhood into adolescence, at least compared to other measures of childhood personality and psychopathology (Frick, Kimonis, Dandreaux, & Farrell, 2003c). There is now fairly substantial evidence that these traits designate an important subgroup of antisocial and delinquent youth (see Frick, 2006; Frick & Marsee, 2006 for other reviews). Frick and Dickens (2006) reviewed 22 published studies showing that CU traits either co-occurred with (n=10), or predicted (n=12), serious antisocial and aggressive behavior, and 5 studies showing that CU traits were related to poorer treatment response among antisocial youth. CU traits are particularly associated with violence that is more premeditated and instrumental in nature (Frick, Cornell, Barry, Bodin, & Dane, 2003a; Kruh, Frick, & Clements, 2005; Pardini, Lochman, & Frick, 2003). Beyond designating a more severe and aggressive subgroup of antisocial youth, CU traits also seem to specify a group of antisocial youth who show characteristics suggestive of different causal processes leading to their antisocial behavior than those operating for other antisocial youth (see Blair, Peschardt, Budhani, Mitchell, & Pine, 2006; Frick & Morris, 2004 for reviews). Specifically, compared to other antisocial youth, youth with CU traits are more likely to show deficits in their processing of negative emotional stimuli (Blair, 1999; Blair, Colledge, Murray, & Mitchell, 2001;

* Corresponding author. Department of Mental Health Law and Policy, Louis de la Parte Mental Health Institute, University of South Florida, 13301 Bruce B. Downs, Tampa, Florida 33612, United States. Tel.: +1 813 974 4510.
E-mail address: kimonis@fmhi.usf.edu (E.R. Kimonis).

0160-2527/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijlp.2008.04.002

Kimonis, Frick, Fazekas, & Loney, 2006; Loney, Frick, Clements, Ellis, & Kerlin, 2003), to show low levels of fearful inhibitions and anxiety (Frick, Cornell, Bodin, Dane, Barry, & Loney, 2003b; Frick, Lilienfeld, Ellis, Loney, & Silverthorn, 1999; Lynam et al., 2005) and to show decreased sensitivity to punishment cues, especially when a reward-oriented response set is primed (Barry et al., 2000; Fisher & Blair, 1998).

These characteristics of antisocial youth with CU traits are important theoretically for at least two reasons. First, they are consistent with developmental theories that have linked problems in conscience development to temperaments characterized by low fearfulness, reward dominance, and lack of emotional responsivity to negative emotional stimuli (Blair, 1995; Frick & Morris, 2004; Kochanska, 1993). Second, similar deficits in performance on laboratory tasks relate to the affective and interpersonal traits of psychopathy during adulthood (Hare & Neumann, 2006; Patrick, Zempolich, & Levenston, 1997). That is, CU traits are one component of the constellation of affective, interpersonal, and behavioral features considered indicative of psychopathy among adults (Cooke & Michie, 1997; Skeem, Mulvey, & Grisso, 2003).

Given this evidence for the importance of CU traits for understanding antisocial and delinquent youth, there is a need for an efficient, reliable and valid measure of these traits. Two of the most widely used measures in research to date (Vincent, 2006) are the PCL-YV (Forth, Kosson, & Hare, 2003) and the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). The PCL:YV has primarily been used in incarcerated samples of adolescents (ages 12 to 18) and utilizes a 60–90 min semi-structured interview and a thorough review of the adolescent's offense records so that highly trained clinicians can rate the youth on 20 items (Vincent, 2006). Four of these items from the PCL-YV are directly related to CU traits.

The APSD is a 20-item rating scale including both parent and teacher (Frick & Hare, 2001), and self-report (Munoz & Frick, 2007) forms, which include 6 items forming the callous–unemotional (CU) subscale. The self-report format has been the most widely used version in adolescent samples and scores from this scale have designated more severe and violent groups of juvenile offenders (Caputo, Frick, & Brodsky, 1999; Kruh et al., 2005), have been associated with an early onset of offending (Silverthorn, Frick, & Reynolds, 2001), and have predicted institutional antisocial behavior and treatment progress in adjudicated adolescents (Spain, Douglas, Poythress, & Epstein, 2004). Although the correlations between the self-report version of the APSD and the PCL-YV have been modest (typically correlations of .30 to .40; Lee, Vincent, Hart, & Corrado, 2003), scores on the APSD have shown comparable correlations with number of arrests (.33) and number of violent arrests (.25) to the PCL-YV (.36 and .28, all $p<.05$) in an adolescent offender sample (Salekin, Leistico, Neumann, DiCicco, & Duros, 2004). Finally, CU traits as measured by the self-report APSD have been associated with deficits in emotional functioning (Kimonis et al., 2006; Loney et al., 2003) and with a lack of sensitivity to punishment in social situations (Pardini et al., 2003) which, as noted previously, are important for causal theories of the development of these traits.

Despite these promising findings for the APSD, it has a number of limitations in its assessment of CU traits. First, only 6 of the 20 items on the APSD measure CU traits and this relatively small number of items probably contribute to its modest internal consistency in many samples (Loney et al., 2003). Further, the small number of items makes it difficult to determine if there are important facets of CU traits that may be differentially related to relevant external criteria (Lynam et al., 2005). Second, items on the APSD are rated on a limited three-point Likert scale with item responses ranging from 0 (Not at all true) to 2 (Definitely true). This limited response format restricts the range and variability of scores. Third, five out of the six items are worded in the same direction, making response sets more likely. Literature on scale construction recommends that questionnaire items include both negatively and positively worded items for a construct (Adkins-Wood, 1961; Anastasi, 1980; Kelloway & Barling, 1990).

To overcome these psychometric limitations of the CU subscale of the APSD, Frick (2004) developed the *Inventory of Callous–Unemotional Traits* (ICU). The development of the ICU involved a number of steps. First, the four items from the APSD CU scale that loaded consistently on the CU factor in clinic and community samples of youth formed the basis for the item content (Frick, Bodin, & Barry, 2000). Second, for each of these four items, three positively- and three negatively-worded items were written to form an item pool of 24 items. These new items, as well as the original APSD items from which they were developed, are presented in Table 1. Third, participants respond to each item based on a 4-point Likert scale that ranges from 0 (*Not at all true*) to 3 (*Definitely true*). Not only does this response format increase the range of responses, but it also does not allow for an exact middle rating.

The first test of the psychometric properties of the ICU was conducted in a large sample ($n=1443$) of 13 to 18 year-old non-referred German adolescents (774 boys and 669 girls; Essau, Sasagawa, & Frick, 2006). Using exploratory factor analysis (EFA), three factors emerged which were labeled Callousness (i.e., "I do not care who I hurt to get what I want"), Uncaring (i.e., "I always try my best", reverse scored), and Unemotional (e.g., "I express my feelings openly", reverse scored). A confirmatory factor analysis (CFA) indicated that a three-factor bifactor model provided acceptable fit to the data ($df=200$, $\chi^2=935.53$, GFI=.90, AGFI=.85, RMSEA=.07). The hallmark of a bifactor model is that in addition to loading on subfactors, all items also load onto a fourth, *general* "callous–unemotional" factor. This type of model has primarily been used in the intelligence literature (e.g., Carroll, 1993; Gustafsson & Balke, 1993), with more recent use in research on adult psychopathy (Patrick, Hicks, Nichol, & Krueger, 2007). This bifactor model fits well for both boys and girls. Also, the scores from the ICU were internally consistent (.77 for the total score) and were correlated with measures of conduct problems, aggression, personality dimensions, and psychosocial impairments in ways that were consistent with past research on CU traits.

Although this initial test of the ICU is promising, there were a number of limitations to this study. First, this best fitting factor model required a large number of correlated error terms ($n=25$) to improve the model fit and such specification can be sample dependent. Thus, this factor structure needs to be replicated in other samples. Second, this sample utilized a German translation of the ICU and this makes it important to determine how stable these findings are across different translations of the scale. Third, this study was limited by its use of a predominantly Caucasian sample. There is a growing body of research suggesting that African American individuals with psychopathic traits do not manifest the same deficits in performance on laboratory measures as

Table 1
Original 24 items on the Inventory of Callous–Unemotional Traits

Careless	Callous
3. **I care about how well I do at school or work (R)**	5. **I feel bad or guilty when I do something wrong (R)**
7. I do not care about being on time	2. What I think is right and wrong is different from what other people think
11. I do not care about doing things well	9. I do not care if I get into trouble
15. I always try my best (R)	13. I easily admit to being wrong (R)
20. I do not like putting the time into doing things well	16. I apologize to persons I hurt (R)
23. I work hard on everything I do (R)	18. I do not feel remorseful when I do something wrong
Unemotional	Uncaring
6. **I do not show my emotions to others**	8. **I am concerned about the feelings of others (R)**
1. I express my feelings openly (R)	4. I do not care who I hurt to get what I want
10. I do not let my feelings control me	12. I seem very cold and uncaring to others
14. It is easy for others to tell how I am feeling (R)	17. I try not to hurt others' feelings (R)
19. I am very expressive and emotional (R)	21. The feelings of others are unimportant to me
22. I hide my feelings from others	24. I do things to make others feel good (R)

Note: The original four CU scale items from the Antisocial Process Screening Device (Frick & Hare, 2001) are in bold print.

Caucasian individuals with such traits (Kosson, Smith, & Newman 1990; Lorenz & Newman, 2002). Fourth, this study involved a sample of non-referred youth. The extent to which the results will characterize referred groups of youthful offenders is unclear.

Thus, in the current study, we explore whether the factor structure identified by Essau et al. (2006) generalizes to a sample of juvenile offenders. In this study, we combined three samples. The first was a sample of boys who had been arrested and were residing in a local detention facility awaiting adjudication; the second was a sample of girls residing in three detention settings serving the same region as the male facility. The third sample included boys who had been arrested, adjudicated, and incarcerated in secure facilities for a sexual offense. In this combined sample, we tested the fit of the factor structure that emerged previously in the German sample and examined the validity of the ICU scales by testing their associations with measures of aggression, delinquency, autonomic reactivity to provocation, and psychosocial functioning.

2. Methods

2.1. Participants

Participants were 248 detained or incarcerated juveniles (188 boys, 60 girls) between the ages of 12 and 20 (M = 15.47, SD = 1.37). The sample was primarily African American (n = 157; 63.3%), with 78 Caucasians (31.5%), 6 Hispanics (2.4%), 2 Native Americans (0.8%), and 4 boys classified as "Other" for ethnicity (1.6%). Four detained boys were excluded from analyses because they had missing data on the ICU scale. As a result, the sample included 98 boys and 60 girls housed in detention facilities, and 90 boys housed in secure confinement facilities following juvenile court disposition for a sexual offense. All facilities were located in or around a large metropolitan area of the Southeastern United States. Table 2 shows the comparison of the three groups on demographic and ICU variables. A series of one-way ANOVAs with sample as the between-groups variable revealed a significant effect of sample for age (F (2,245) = 7.03, p < .001), the ICU Callousness factor (F (2,245) = 4.59, p < .05), the ICU Unemotional factor (F (2,245) = 6.49, p < .01), and the ICU total score (F (2,245) = 5.01, p < . 01). Overall, the sample of girls tended to be younger and scored lower on the Unemotional dimension, whereas sex offenders tended to score lowest on ICU total, Uncaring, and Callousness scores.

Table 2
Characteristics of the three samples

Variable	Detained boys (n = 98)	Detained girls (n = 60)	Male sex offenders (n = 90)	Full sample (n = 248)
Age [a]	15.50 (1.26) [a]	14.95 (1.29) [b]	15.79 (1.45) [a]	15.47 (1.37)
Ethnicity	69.4/ 21.4	78.3/ 21.7	46.7/ 48.9	63.3/ 31.5
ICU Uncaring	9.28 (4.93) [a]	9.12 (5.01) [a, b]	7.73 (5.42) [b]	8.68 (5.16)
ICU Callousness [b]	6.21 (4.49) [a]	5.50 (4.22) [a, b]	4.13 (5.32) [b]	5.29 (4.82)
ICU Unemotional [c]	8.08 (2.94) [a]	6.35 (3.06) [b]	7.64 (2.94) [a]	7.50 (3.03)
ICU Total score [d]	26.07 (8.25) [a]	23.73 (9.23) [a, b]	21.80 (10.27) [b]	23.96 (9.41)

Note: Effects are from a one-way ANOVA with sample as the between-groups factors. Different letters denote significant differences in pairwise comparisons using the LSD procedure for pairwise comparisons; Numbers in ethnicity cells indicate the percentage of African Americans/Caucasians.

[a] F (2, 245) = 7.03, p < .001, Partial Eta^2 = .05.
[b] F (2, 245) = 4.59, p < .05, Partial Eta^2 = .04.
[c] F (2, 245) = 6.49, p < .01, Partial Eta^2 = .05.
[d] F (2, 245) = 5.01, p < .01, Partial Eta^2 = .04.

2.2. Procedures

For the two detained samples, a staff member from each detention center contacted the parents or legal guardians of all youth currently residing at the facility and informed them of a study being conducted by researchers at a local university and asked permission to forward their phone number to the researchers. They were informed that their child's participation in the project would in no way influence his or her treatment at the detention center or his or her legal standing in the adjudication process. Those parents who agreed to be contacted by the researchers were phoned and had the study procedures explained to them. As approved by the host university's Institutional Review Board and the director of the detention center, parents or legal guardians who agreed to have their child participate were asked to have the consent process tape-recorded and were subsequently mailed a copy of the consent form for their records. Youth whose parents provided consent were met in a private room at the detention center and were asked to assent to participate. Of those youth whose parents were contacted, 81% of detained boys and 73% of detained girls participated in the study. For all male participants who had parental consent and child assent, the provocation task described below was administered individually during which psychophysiological indices of reactivity were collected. For both male and female samples, all self-report measures were administered in small groups (3 to 8 participants) at the detention centers and all questionnaires were read aloud to control for reading level. Following completion of the questionnaires, each participant received either a soft drink and candy bar (male sample) or pizza (female sample).

For the sex offender sample, participant information was obtained from an electronic extraction of case record information. This information was a subset of data from a broader archival study of intake admission and assessment records from youth in the secure custody institution. All records were extracted without identifying information. Due to the archival nature of this project, and confidentiality protections built into the record extraction process, the Institutional Review Board waived informed consent requirements. As a standard part of the facility assessment process for sexually offending youth, all youth with a current sexual offense were administered the ICU and other specialized assessment instruments including the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002). Psychology staff administered the ICU as part of the overall assessment protocol. Working collaboratively with psychology staff, social work staff rated YLS/CMI items following a standardized interview with the youth, collateral phone interviews with parents/legal guardians, and a review of all available case record materials. All assessment instruments were completed within the first 30 days of admission to the facility.

2.3. Measures

2.3.1. Inventory of Callous–Unemotional Traits (ICU; Frick, 2004)

The ICU was administered to all three samples. It includes 24 items (e.g., "I do not show my emotions to others") that are rated on a four-point Likert scale from 0 (Not at all true) to 3 (Definitely true). A thorough description of the creation of this measure is provided in the Introduction and information on its reliability and validity in this sample are reported in the Results section.

2.3.2. The Antisocial Process Screening Device (APSD: Frick & Hare, 2001)

The APSD was administered at the male detention center only. This self-report measure includes 20 items that are rated on a three-point scale from 0 (Not at all true) to 2 (Definitely true). The Callous–Unemotional (CU) subscale of the APSD, which formed the basis for the ICU item content, was used in the current study. A description of this measure and past evidence for its validity is provided in the Introduction. The CU scale demonstrated adequate internal consistency (α=.71) in the current sample.

2.3.3. Peer Conflict Scale (PCS; Kimonis, Marsee, & Frick, 2004)

The PCS was administered as a self-report measure of aggression to the two detained samples only. It was developed to improve upon existing measures for assessing aggression by a) measuring four dimensions of aggression (i.e., reactive overt, proactive overt, reactive relational, proactive relational), b) including a sufficient number of items (n=10) for each dimension, and c) limiting items to only acts clearly harming another person. Items were pooled from a number of aggression scales (Bjorkqvist, Lagerspetz, & Osterman, 1992; Brown, Atkins, Osborne, & Milnamow, 1996; Crick & Grotpeter, 1995; Dodge & Coie, 1987; Galen & Underwood, 1997; Little, Jones, Henrich, & Hawley, 2003). Redundant items and items that weren't clearly related to harming others were deleted. Items were reworded to ensure that there was direct correspondence between overt and relational items, such that for each reactive overt item (e.g., "I hurt others when I am angry at them") there was an analogous reactive relational item (e.g., "Sometimes I gossip about others when I'm angry at them"), and for each proactive overt item (e.g., "I start fights to get what I want") there was an analogous proactive relational item (e.g., "I try to make others look bad to get what I want"). These items were then reviewed by a team of faculty, graduate, and advanced undergraduate students to ensure that the wording was developmentally appropriate. Items are rated on a 4-point Likert scale from 0 ("Not at all true") to 3 ("Definitely true"). All four subscales demonstrated adequate internal consistency in this sample, ranging from .77 (proactive overt) to .87 (reactive overt).

2.3.4. Self-Reported Delinquency Scale (SRD; Elliot & Ageton, 1980)

The SRD scale was also administered to both detained samples. It assesses the types of crimes committed by the youth. The SRD lists 36 questions about illegal juvenile acts selected from a list of all offenses reported in the Uniform Crime Report with a juvenile base rate of greater than 1% (Elliott & Huizinga, 1984). For each question the youth is asked to respond with a "yes" or "no" regarding whether he/she has ever done the behavior. Consistent with past uses of the scale (Krueger et al., 1994), a total delinquency composite was created by summing the number of delinquent acts committed (with a possible range of 0–36). In

addition to the total score, the current study also used the 20-item non-violent offenses subscale (e.g. property, drug, and status offenses) and the 8-item violent offenses subscale (e.g., "have you ever been involved in gang fights?"). All subscales demonstrated adequate internal consistency in this sample, ranging from .61 (violent delinquency) to .88 (total delinquency).

2.3.5. BarOn Emotion Quotient Inventory (EQI; Bar-On & Parker, 2000)

The EQI was administered at the male detention center only. It is a self-report measure that was used to assess socioemotional competence. This study included a 5-item Empathy scale (e.g., "I feel bad when other people have their feelings hurt") that was created using items from the Intrapersonal scale of the EQI, and a 13-item Positive Affect scale (e.g., "I am happy"; "I know things will be okay"), which was created using items from the General Mood scale of the EQI. Items are rated on a four-point Likert scale, ranging from "Agree Strongly" to "Disagree Strongly," with higher scores indicating better socioemotional competence. Past research has supported the construct validity of these scales by showing expected convergent and divergent correlations with the factors of the NEO-Five Factor Inventory, depressive symptomatology, and externalizing and internalizing problematic behaviors (Bar-On & Parker, 2000). The two subscales demonstrated adequate internal consistency in this sample, ranging from .65 (Empathy) to .85 (Positive Affect).

2.3.6. Autonomic reactivity

Only youth at the detention center for boys completed a computerized provocation task, the *Competitive Reaction Time Task* (CRTT; Waschbusch et al., 2002), that included three levels of provocation from a fictitious peer. Each child was seated in front of a computer screen and read an instructional script, informing them that they would be playing a computer game against a boy from another facility. On each trial, a target appeared on the screen to which the participant was to press the space bar as fast as possible. If they responded faster than their fictitious opponent, they would earn 50 points and they could take 0 to 100 points in increments of 10 from their opponent. For each participant, the game was pre-programmed for the same 16 losses out of 48 trials. Two losing trials never occurred in succession. Standard pre-recorded verbal messages by a young African American male from the local community were played over the computer when a loss occurred. Eight of 16 loss trials were high provocation trials, whereby a highly aversive verbal message was played (e.g., "You wimp! I don't think I'll ever be beaten! Minus 100!") and 80–100 points were subtracted by the opponent. The other eight of the 16 loss trials were low provocation trials, whereby a less-provoking verbal message (e.g., "I won, but I'll give you a break. I'll only take 10 points") was broadcast and 0–20 points were subtracted. The computer indicated a win on the remaining 32 of the 48 trials, resulting in a net win of 780 points. After completion of the computer game, youth completed a questionnaire to determine whether the deception of the hypothetical peer was successful. Also, after the participant was released from the detention center, a letter thanking them for their participation and debriefing them about the deception used for the provocation was sent to the participant's home. This debriefing was done following release from the center to avoid the participants sharing this information with other potential participants in the facility.

Three participants expressed some doubts about the legitimacy of the task but were retained in the sample because eliminating them did not influence the results. Separate aggressive response measures were computed based on the level of provocation. A measure of aggressive responding to no provocation was obtained by examining aggressive responding during the first three win trials and before experiencing a provocation. In addition, aggressive responding was averaged for the trials immediately following low and high provocation trials. Supporting this manipulation, participants responded with more aggression following high provocation ($M=86.79$; $SD=18.69$) compared to low provocation ($M=65.21$; $SD=28.30$) ($F(1, 96)=82.14$; $p<.001$; partial $eta^2=.46$) and with more aggression following low compared to no provocation ($M=55.50$; $SD=38.42$) ($F(1, 96)=79.43$; $p<.001$; partial $eta^2=.45$).

During the CRTT, measures of autonomic reactivity to the two levels of provocation were recorded. Electrodermal activity (EDA) for determining skin conductance level (SCL) was recorded via two electrodes placed on the middle two distal phalanges of the non-dominant hand using Thought Technology's ProComp Infinity encoder connected to a Pentium 4 laptop computer equipped with Biograph Infinity software (version 2.0.1). Sampling for EDA was set at 256 Hz. After a 10-minute stabilization period, autonomic activity was measured for 3 min prior to the CRTT (baseline period) and during the 9- to 11-minute CRTT. Separate skin conductance response (SCR) scores were determined for periods following low and high provocation. After the end of each taunt, the change (0.01 microsiemens or greater) in SCL between the 1-second and 4-second mark was obtained and averaged for skin conductance response (SCR) to low and high provocation (Stern et al., 2001).

2.3.7. Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002)

The YLS/CMI was completed for all participants in the sex offender sample. This inventory is a standardized checklist of risk/needs factors that are used to classify youth's individual and overall levels of risk. Social work staff worked collaboratively with psychology staff in completing the intake mental health assessments. Following a standardized interview with the youth, collateral phone contact with the parent/legal guardian, and record review, social work staff assigned to the case rated the YLS/CMI.

The YLS/CMI assesses eight different risk/needs areas: Prior and Current Offenses/Disposition (e.g., number of prior and current convictions, failures to comply), Family Circumstances/Parenting (e.g., inadequate supervision, parental difficulty in controlling the behavior of the youth), Education/ Employment (e.g., disruptive school behavior, negative relationships with teachers and school peers), Peer Relations (e.g., absence of positive acquaintances/friends, association with delinquent acquaintances/friends), Substance Abuse (e.g., various levels of substance use, a connection between substance use and offending behavior), Leisure/Recreation, Personality/Behavior, and Attitudes/Orientation (e.g., antisocial/ procriminal attitudes, callousness, active rejection of help). For the current analyses the total risk score summing all eight risk/need scores (alpha=.89), the Prior Offenses/Disposition

score, and the Attitudes/Orientation score were used in analyses to validate the ICU. The validity of these scores is supported by past research showing that the YLS/CMI total risk score is significantly correlated with indices of reoffending, externalizing disorders, and the callous/deceitful and conduct problems factor scores of an early version of the Hare Psychopathy Checklist—Youth Version (Forth et al., 2003; see Hoge, 2005).

3. Results

3.1. Confirmatory factor analysis

AMOS 5.0 (Arbuckle, 2003) with maximum likelihood estimation was used for all confirmatory factor analyses (CFAs). Three a priori models were tested. First, we tested a unidimensional model. This was tested as a baseline model to which we could compare other factor structures. Table 3 provides the fit statistics for this and other factor models that were estimated. Model fit was evaluated using the χ^2 fit statistic, χ^2/df ratio, comparative fit index (CFI; Bentler, 1990), and the root mean square error of approximation (RMSEA). A good fit was determined by values of the Comparative Fit Index (CFI; Bentler, 1990) above .95. A Root Mean Square Error of Approximation (RMSEA) below .05 indicates a very good model, a value of .08 indicates adequate fit, and a value above .10 indicates poor model fit (Browne & Cudeck, 1993). According to these criteria, the results of this analysis revealed an inadequate fit to the data for the unidimensional model (df=252, χ^2=890.76, p<.001, χ^2/df=3.53, CFI=.50, RMSEA=.10).

Because the ICU scale was originally developed from four items on the APSD CU scale (Table 1), we next examined a four-factor hierarchical model that reflected the scale's development strategy. In this model, the 24 ICU items loaded separately on the four lower-order factors that reflected the original APSD CU scale items. In turn, these four factors loaded on a single higher-order "callous–unemotional" factor. The results of this analysis revealed inadequate fit to the data (df=249, χ^2=800.13, p<.001, χ^2/df=3.21, CFI=.57, RMSEA=.10).

Next, we tested the fit of the three-factor bifactor model found in a large sample of community German adolescents by Essau et al. (2006) but without the 25 correlated error terms. The bifactor model is an alternative approach that has important advantages over the hierarchical/second-order model (see Chen, West, & Sousa, 2006). The hierarchical approach represents subfactors as correlated components of a higher-order construct. In contrast, the bifactor model specifies that there is a general "callous–unemotional" factor that underlies each of the items (reflecting the overlap across all items of the ICU). Separately, there are also three independent subfactors (i.e., Callousness, Uncaring, and Unemotional), each of which account for unique variance in their respective set of items, over and above the variance accounted for by the general factor (see Fig. 1). Further, given that the subfactors are related to the contribution that is over and above the general factor, they are assumed to be uncorrelated. Although this three-factor bifactor model fit significantly better than the previously estimated unidimensional (df=21, χ^2=388.46, p<.001) and four-factor hierarchical (df=18, χ^2=297.83, p<.001) models, its absolute level of fit was still unacceptable (df=231, χ^2=502.30, p<.001, χ^2/df=2.17, CFI=.79, RMSEA=.07).

Having found inadequate fit for all a priori models tested, we (a) examined the data to determine whether particular items contributed to poor fit, (b) deleted those items, and (c) conducted CFAs that fitted two promising three-factor models. Item-total correlations indicated that items 2 (r=−.01; "What I think is right and wrong is different from what other people think") and 10 (r=.04; "I do not let my feelings control me") from the Callousness dimension were essentially unrelated to the remaining items on the scale. These items also manifested low factor loadings in Essau et al.'s (2006) previous factor analyses with a German Sample. Therefore, these items were deleted. With these items deleted, we completed CFAs for a three-factor hierarchical model and Essau et al.'s alternative three-factor bifactor model (The four-factor model was not completed, given deletion of items that comprised one of the original factors). Results indicate that the three-factor hierarchical model inadequately fit the data (df=206, χ^2=471.25, p<.001, χ^2/df=2.29, CFI=.79, RMSEA=.07). However, the three-factor bifactor model showed inadequate fit by one index (CFI) and adequate fit with another (RMSEA) (df=187, χ^2=343.52, p<.001, χ^2/df=1.84, CFI=.87, RMSEA=.06). The revised bifactor model fit significantly better than the three-factor hierarchical model with the two poorly loading items deleted (df=19, χ^2=127.73, p<.001). The model specification for this bifactor model is presented in Fig. 1.

In Table 4, factor loadings for the three-factor bifactor model are presented alongside loadings for the three-factor hierarchical model. The magnitude of factor loadings for the three-factor hierarchical model demonstrate that the Callousness and Uncaring factors are well-represented by their respective items, although the Unemotional scale includes some low-loading items. The loadings from the bifactor model show the variance accounted for by the general factor in addition to the unique variance, over and above that accounted for by the general factor, that is accounted for by each subfactor, Callousness, Uncaring, and Unemotional. As

Table 3
Fit indices comparing the confirmatory factor models for the Inventory of Callous–Unemotional Traits (ICU)

Model	Chi-Sq (χ^2)	df	CFI	RMSEA	χ^2/df
1. Unidimensional Model	890.76*	252	0.50	0.10	3.53
2. Original 4-Factor Hierarchical Model	800.13*	249	0.57	0.10	3.21
3. German 3-Factor Bifactor Model (without correlated errors)	502.30*	231	0.79	0.07	2.17
4. 3-Factor Hierarchical Model (without items 2 and 10)	471.25*	206	0.79	0.07	2.29
5. 3-Factor Bifactor Model (without items 2 and 10)	343.52*	187	0.87	0.06	1.84

Note: *p<.001; CFI—Comparative fit index (CFI; Bentler, 1990); RMSEA = root mean square error of approximation (Bentler, 1995; Ullman, 1996).

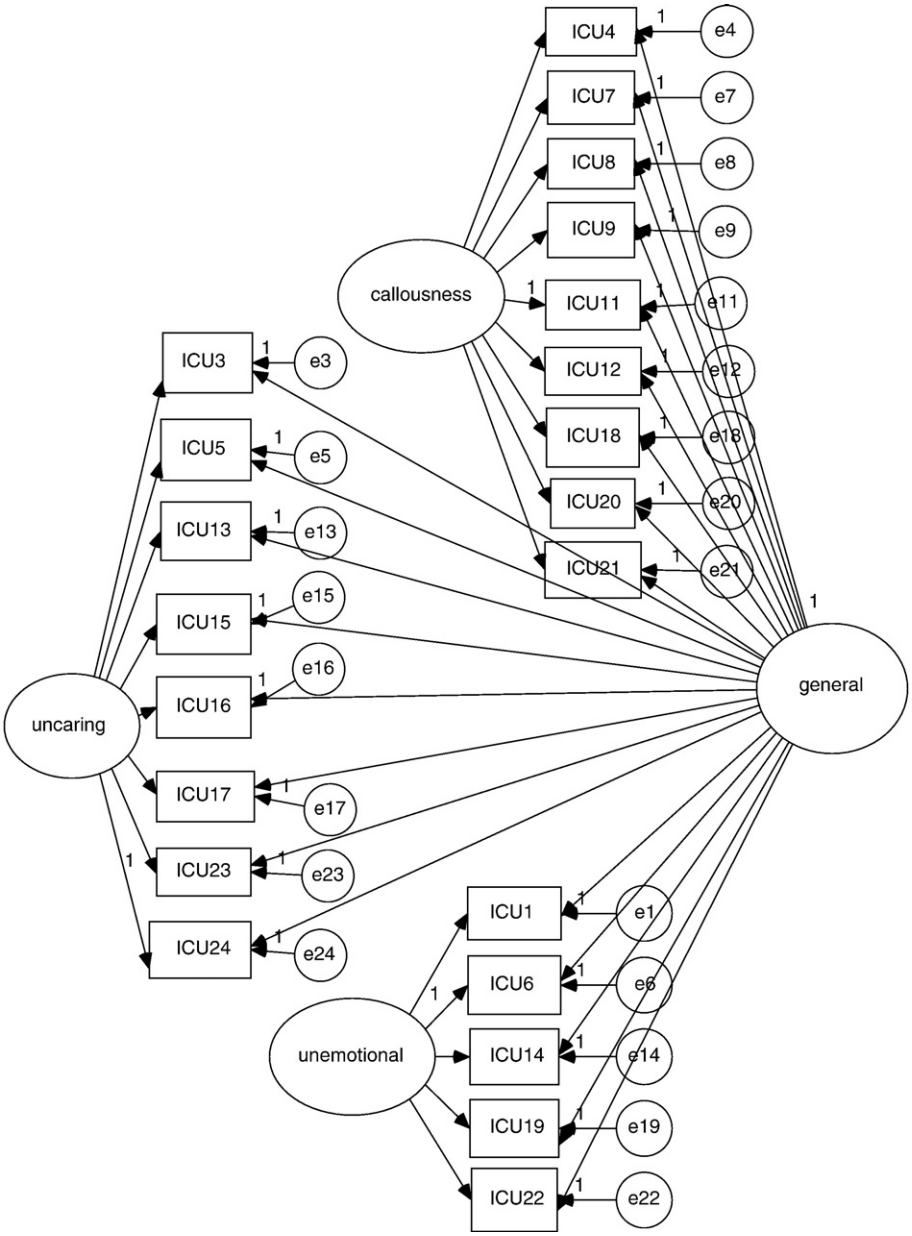

Fig. 1. Three-factor bifactor structural model of the general and specific factors of the Inventory of Callous–Unemotional Traits.

illustrated by these loadings provided in Table 4, items on the Uncaring factor showed the strongest loadings on the general factor (loadings ranging from −.41 to −.76 all $p<.05$). Thus, after controlling for this general factor, the factor loadings specific to the Uncaring factor tended to be low. For example, the item "I apologize to persons I hurt" showed the highest loading on the general factor (−.76) but, after controlling for this general factor in the bifactor model, it was only modestly associated with the Uncaring factor (−.15). Interestingly, the items from the Callousness and Unemotional factors had relatively limited loadings on the general factor.

3.2. Internal consistency

Based on this factor analysis, four scales were created by inversely scoring items to all be in the callous–unemotional direction, summing items, and eliminating items 2 and 10. The coefficient alpha for the Total ICU scale combining all 22 items in the combined sample was .81, and for the three subscales were .81, .80, and .53 for Uncaring, Callousness, and Unemotional, respectively. For the Unemotional scale, inspection of the item-total correlations did not suggest that the deletion of any single item would significantly improve the internal consistency of the scale. The small number of items ($n=5$) constituting this subscale may

Table 4
Factor loadings for the best fitting three-factor bifactor model (shaded) and the three-factor hierarchical model for the Inventory of Callous–Unemotional Traits (ICU)

Items	G	Uncaring		Callousness		Unemotional	
Model	BF	BF	H	BF	H	BF	H
Uncaring							
*23. I work hard on everything I do.	−.53	.71	.73				
*15. I always try my best.	−.52	.48	.69				
*3. I care about how well I do at school or work.	−.41	.38	.55				
*24. I do things to make others feel good.	−.52	.22	.55				
*16. I apologize ('say I am sorry') to persons I hurt.	−.76	−.15^{ns}	.58				
*5. I feel bad or guilty when I do something wrong.	−.54	.15^{ns}	.57				
*13. I easily admit to being wrong.	−.50	.03^{ns}	.44				
*17. I try not to hurt others' feelings.	−.70	.01^{ns}	.62				
Callousness							
11. I do not care about doing things well.	.16			.65	.66		
20. I do not like to put the time into doing things well.	.09^{ns}			.60	.59		
18. I do not feel remorseful when I do something wrong.	.19			.58	.62		
7. I do not care about being on time.	.04^{ns}			.56	.54		
9. I do not care if I get into trouble.	.20			.54	.57		
12. I seem very cold and uncaring to others.	.14			.52	.54		
21. The feelings of others are unimportant to me.	.12^{ns}			.50	.52		
4. I do not care who I hurt to get what I want.	.29			.42	.49		
*8. I am concerned about the feelings of others.	−.41			−.32	−.44		
Unemotional							
6. I do not show my emotions to others.	−.02^{ns}					.56	−.48
*1. I express my feelings openly.	−.19					−.53	.62
22. I hide my feelings from others.	−.02^{ns}					.48	−.46
*14. It is easy for others to tell how I am feeling.	−.22					−.27	.28
*19. I am very expressive and emotional.	−.36					−.22	.29

Note: G = General Factor; BF = Bifactor model; H = Hierarchical model; * = Reverse scored items. N=248. All factor loadings are significant at $p<.05$, except where denoted by ^{ns}. All factors are uncorrelated.

explain the low internal consistency found. The coefficient alphas were consistent across samples with the internal consistency for the Total ICU score ranging from .74 to .85 across samples. The alphas ranged from .78 to .84 for the Uncaring scale, from .71 to .88 for the Callousness scale, and from .45 to .60 for the Unemotional scale. The subscales were weakly correlated with one another with correlations of .29 ($p<.001$) and .23 ($p<.001$) between Uncaring and Callousness, and Uncaring and Unemotional, respectively, and .17 ($p<.01$) between Callousness and Unemotional.

3.3. Construct validity

Not all indices used to test the construct validity of the ICU and the component scales were present in all three samples. The first associations tested were between the ICU and measures of self-reported aggression and self-reported delinquency in the detained samples of boys and girls. These correlations are reported in Table 5. The Total ICU scale was generally associated with all four types of aggression (proactive overt, reactive overt, proactive relational, reactive relational) and all three measures of self-reported delinquency (total, violent, non-violent). Across delinquency types, 13 of the 14 correlations were statistically significant ($p<.05$) and ranged from r=.16 to r=.44. Only one correlation with violent delinquency (in detained boys) was below .25. Also evident from the correlations reported in Table 5 was that the Unemotional dimension was not strongly related to the measures of aggression and delinquency, showing only one significant correlation (r=.26 with reactive overt aggression in detained girls). Further, the Callousness dimension showed more consistent correlations with aggression, whereas the Uncaring dimension seemed to be more strongly and consistently associated with the delinquency measures.

For the sample of detained boys, correlations were computed between the ICU scale and the Empathy and Positive Affect scales of the EQI. Again, the ICU Total score was associated with both of these measures of emotional functioning (r=−.51 and r=−46,

Table 5
Correlations between Inventory of Callous–Unemotional Traits (ICU) factors and external criteria

Variable	Total ICU	Uncaring	Callousness	Unemotional	APSD CU
Combined detained samples					
AGGRESSION					
Proactive overt					
Detained boys (n=98)	.37***	.34***	.25*	−.15	.22*
Detained girls (n=60)	.41***	.18	.52***	.05	
Reactive overt					
Detained boys	.27**	.19[a]	.16	.06	.22*
Detained girls	.30*	.07	.34**	.26*	
Proactive relational					
Detained boys	.36***	.29**	.24*	−.08	.16
Detained girls	.44***	.11	.50***	.21	
Reactive relational					
Detained boys	.28**	.12	.23*	.02	.12
Detained girls	.42***	.10	.43***	.24[a]	
DELINQUENCY					
Total					
Detained boys	.26*	.33***	.04	−.04	.03
Detained girls	.38**	.33**	.21	.10	
Violent					
Detained boys	.16	.19[a]	.06	−.07	−.01
Detained girls	.39**	.17	.45***	.06	
Non-violent					
Detained boys	.26**	.33***	.04	−.04	.05
Detained girls	.34**	.34**	.11	.13	
Detained boys					
SOCIOEMOTIONAL					
Empathy	−.51***	−.46***	−.05	−.33***	−.28**
Positive affect	−.38***	−.21*	−.16	−.28**	−.40***
PSYCHOPHYSIOLOGY					
Mean SCR high provocation	−.20*	−.20*	−.04	−.10	−.19[a]
Mean SCR low provocation	−.21*	−.12	−.12	−.11	−.14
Sex offenders(n=90)					
Previous offenses /dispositions	.27**	.34***	.04	.12	
Attitudes/orientation	.27**	.43***	.06	−.03	
YLS total risk	.33***	.44***	.13	.01	

Note: APSD CU = Antisocial Process Screening Device (Frick & Hare, 2001) Callous–Unemotional Scale; YLS = Youth Level of Service/Case Management Inventory (Hoge & Andrews, 1994, 2002). *p<.05;**p<.01;***p<.001; [a]p=.06.

p<.001), indicating that the ICU was associated with less empathy and less positive affect on the EQI. The Uncaring scale was also negatively associated with these measures of emotional functioning (r=−46, p<.001 and r=−.21, p<.05). In contrast to the findings for aggression and delinquency, the Unemotional dimension was also associated with both indices from the EQI (r=−.33, p<.001 and r=−.28, p<.01).

In Table 5, the correlations between the ICU and measures of skin conductance reactivity at both high and low levels of provocation during the computer task are also reported. Again, this task was only conducted with the detained boys. The total score was negatively related to measures of reactivity to provocation at both high (r=−.20, p<.05) and low (r=−.21, p<.05) levels of provocation. These validity coefficients are much lower than those reported for the aggression, delinquency, and emotional functioning measures. However, given that the former measures were all self-report, shared method variance with the self-reported ICU could have inflated these correlations. Also, the only correlation with skin conductance reactivity to reach significance for the ICU subscales was between the Uncaring scale and reactivity to high provocation (r=−.20, p<.05).

In the sample of male sex offenders, the ICU scales were correlated with the Previous Offenses/Dispositions, and the Attitudes/Orientation subscales, as well as the total risk scale from the YLS/CMI. Scoring of this measure includes self-report by the youth but also includes reports from parents and information from the youth's records. As evident from Table 5, correlations with these scores from the YLS/CMI were similar to the correlations found for the self-report of delinquency. That is, the Total ICU scale was correlated with previous offenses (r=.27, p<.01), procriminal attitudes (r=.27, p<.01), and overall risk for offending (r=.33, p<.001), with the Uncaring subscale seeming to account for most of this association (r=.34, .43, and .44, p<.001).

To compare the ICU and its subscales with the CU scale of the APSD we examined the correlations between them in the detained sample of boys. The total score of the ICU and its subscales were moderately correlated with the CU scale of the APSD (r=.45, ICU total; r=.32, Uncaring; r=.36 Callous; all p<.001), with the exception of the Unemotional subscale of the ICU (r=.14, n.s.). Correlations between the APSD CU scale and indices used to test the construct validity of the ICU scale are presented alongside the ICU subscales in Table 5. The APSD CU scale and the ICU total scale showed similar associations with validity variables. However, the ICU total score was significantly associated with validity variables in 20 of 21 correlations, whereas correlations reached

significance for the APSD CU scale in only 4 of 11 correlations. The stronger associations were particularly evident in the correlations between the ICU total scale and measures of self-reported delinquency.

4. Discussion

The current study is the first to explore the psychometric properties of the English version of the *Inventory of Callous–Unemotional Traits* scale (ICU; Frick, 2004) in a sample of adolescent offenders (n=248). The results suggest that this scale is promising as an extended measure of CU traits relative to the PCL:YV (Forth et al., 2003) and APSD (Frick & Hare, 2001). First, the bifactor confirmatory analysis is consistent with a general factor present across the ICU items. Further, the total score from the ICU was moderately correlated with the six-item CU scale from the APSD, but showed much improved internal consistency in comparison (Loney et al., 2003). Also, the validity coefficients reported in Table 5 suggest that the total ICU score was correlated with self-reported measures of aggression and delinquency, with both self-reported and psychophysiological indices of constricted emotion, and with measures of past offending that included reviews of institutional records and collateral reports, all of which have been important correlates to CU traits in past studies (see Frick, 2006; Frick & Marsee, 2006 for reviews). Thus, the ICU total score has proven to show validity in a community sample of German Caucasian adolescents (Essau et al., 2006) and in this ethnically diverse sample of detained adolescents from the United States.

Second, these results broadly are consistent with the factor structure obtained previously in the German sample, even without using correlated error terms that can be sample dependent. That is, confirmatory factor analyses suggested that the total ICU scale consists of three relatively independent dimensions of behavior: Uncaring (e.g., "I work hard on everything I do"—reverse scored), Callousness (e.g., "I do not care about doing things well"), and Unemotional ("I express my feelings openly"—reverse scored). The fact that similar factors emerge in two such diverse samples of adolescents and using two different languages supports this factor structure.

Third, the differential correlations of ICU factors with external correlates provide some preliminary data to suggest that different facets of the callous–unemotional dimension may show specific associations with some of the correlates to CU traits that have been documented in past research. To summarize the correlations reported in Table 5, the Callousness dimension seemed to be more strongly associated with the measures of aggression, whereas the Uncaring dimension seemed to be more strongly related to measures of offending. In contrast, the associations with the Unemotional dimension were specific to the measures of emotional functioning (i.e., lack of empathy; lack of positive affect). These findings obviously need to be replicated in different samples, and using other measures. However, they suggest that CU traits, which appear to be very important for understanding antisocial youth (Frick & Dickens, 2006), may be a constellation of several related facets of affective and interpersonal functioning that may each be distinctly related to specific impairments and could potentially have distinct causal factors (Lynam et al., 2005).

These results must be considered in light of several study weaknesses. First, for the three-factor model to approach adequate fit to the data, two items from the ICU had to be deleted. Thus, further testing of the item set is required to see if this finding is sample dependent, although the results from Essau et al. (2006) also raised concerns about these items. Another concern was the low internal consistency of the Unemotional factor, which may be explained by its small number of constituent items (n=5). The internal consistency and construct validity of this subscale needs to be tested in other samples of youth before firm conclusions about its utility can be made. Second, whereas the total score of the ICU is made up of equal numbers of positively and negatively worded items, two of the subscales that emerged consisted largely of negatively worded (Callousness) or positively worded (Uncaring) items. Thus, it is possible that method variance, and not construct variance, contributed to the grouping of items (Burke, 1999; Cordery & Sevastos, 1993; Schmitt & Stults, 1985). Third, and also related to method variance, we did not have another measure of CU traits assessed through a different method, such as the PCL-YV, to determine how strongly the measures correlated. This is a critical issue given that psychopathy measures in general have been shown to have strong method variance, with measures using similar assessment formats showing substantial correlations, but with correlations across formats being quite modest (Lee et al., 2003). Fourth, the study combined three different samples of offending youth and the size of the individual samples did not allow for a test of factor invariance across groups. Another limitation of the sample was the wide age range of the youth, which spanned from 12 to 20.

Thus, all interpretations need to be made in the context of these limitations. However, there were also a number of important strengths to the current study. Correlates to the ICU were assessed with multiple methods (e.g., self-report, psychophysiology, clinical interview and collateral reporters) and the samples included substantial numbers of ethnic minority adolescents and included both boys and girls. As a result, the findings provide indications that the ICU is promising as a measure of a construct, callous–unemotional traits, that has proven to be very important for designating a distinct subgroup of antisocial youth. As noted in the Introduction, CU traits constitute only one dimension of the construct of psychopathy. However, some have argued that it may be one of the most important dimensions to this personality disturbance, especially for differentiating within antisocial individuals (Barry et al., 2000; Skeem & Cooke, submitted for publication). Therefore, not only may these traits be important for understanding a group of youth who show very severe and aggressive antisocial behaviors, it may be critical for understanding the developmental precursors to a very serious form of personality disturbance.

References

Adkins-Wood, D. (1961). *Test construction.* Columbus OH: Merrill.
Anastasi, A. (1980). *Psychological testing.* New York: MacMillan.

Arbuckle, J. L. (2003). *Amos Version 5.0.* Chicago: Small Waters.
Bar-On, R., & Parker, J. D. A. (2000). *BarOn Emotional Quotient Inventory: Youth version technical manual.* North Tonawanda, NY: Multi-Health Systems.
Barry, C. T., Frick, P. J., Grooms, T., McCoy, M. G., Ellis, M. L., & Loney, B. R. (2000). The importance of callous–unemotional traits for extending the concept of psychopathic traits to children. *Journal of Abnormal Psychology*, *109*, 335–340.
Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238–246.
Bentler, P. M. (1995). *EQS structural equation program model.* Encino, CA: Multivariate Software.
Bjorkqvist, K., Lagerspetz, K. M. J., & Osterman, K. (1992). *The direct and indirect aggression scales.* Vasa, Finland: Abo Akademi University, Department of Social Sciences.
Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, *57*, 1–29.
Blair, R. J. R. (1999). Responsiveness to distress cues in the child with psychopathic tendencies. *Personality and Individual Differences*, *27*, 135–145.
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*(4), 491–498.
Blair, R. J. R., Peschardt, K. S., Budhani, S., Mitchell, D. G. V., & Pine, D. S. (2006). The development of psychopathy. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 262–275.
Brown, K., Atkins, M. S., Osborne, M. L., & Milnamow, M. (1996). A revised teacher rating scale for reactive and proactive aggression. *Journal of Abnormal Child Psychology*, *24*(4), 473–480.
Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
Burke, B. G. (1999). Item reversals and response validity in the job diagnostic survey. *Psychological Reports*, *85*, 213–219.
Caputo, A. A., Frick, P. J., & Brodsky, S. L. (1999). Family violence and juvenile sex offending: The potential mediating role of psychopathic traits and negative attitudes toward women. *Criminal Justice & Behavior*, *26*(3), 338–356.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge: Cambridge University Press.
Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–225.
Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist—Revised. *Psychological Assessment*, *9*(1), 3–14.
Cordery, J. L., & Sevastos, P. P. (1993). Responses to the original and revised job diagnostic survey: Is education a factor in responses to negatively worded items? *Journal of Applied Psychology*, *78*(1), 141–143.
Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-aggression among children. *Developmental Psychology*, *33*, 589–600.
Dodge, K. A., & Coie, J. D. (1987). Social-information processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, *52*, 1146–1158.
Elliott, D. S., & Ageton, S. (1980). Reconciling ethnicity and class differences in self-reported and official estimates of delinquency. *American Sociological Review*, *45*(1), 95–110.
Elliott, D. S., & Huizinga, D. (1984). *The relationship between delinquent behavior and ADM problems.* Boulder, CO: Behavioral Research Institute.
Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous–unemotional traits in community sample of adolescents. *Assessment*, *13*, 454–469.
Fisher, L., & Blair, R. J. R. (1998). Cognitive impairment and its relationship to psychopathic tendencies in children with emotional and behavioral difficulties. *Journal of Abnormal Child Psychology*, *26*, 511–519.
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. Unpublished rating scale.
Frick, P. J. (2006). Developmental pathways to conduct disorder. *Child and Adolescent Psychiatric Clinics of North America*, *15*(2), 311–331.
Frick, P. J., & Dickens, C. (2006). Current perspectives on conduct disorder. *Current Psychiatry Reports*, 59–72.
Frick, P. J., & Hare, R. D. (2001). *The Antisocial Process Screening Device (APSD).* Toronto: Multi-Health Systems.
Frick, P. J., & Marsee, M. A. (2006). Psychopathic traits and developmental pathways to antisocial behavior in youth. In C. J. Patrick (Ed.), *Handbook of psychopathic traits* (pp. 355–374). New York: Guilford.
Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *33*(1), 54–68.
Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the psychopathy screening device. *Psychological Assessment*, *12*(4), 382–393.
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.
Frick, P. J., Cornell, A. H., Bodin, S. D., Dane, H. E., Barry, C. T., & Loney, B. R. (2003). Callous–unemotional traits and developmental pathways to severe conduct problems. *Developmental Psychology*, *39*(2), 246–260.
Frick, P. J., Kimonis, E. R., Dandreaux, D. M., & Farrell, J. M. (2003). The 4-year stability of psychopathic traits in non-referred youth. *Behavioral Sciences & the Law*, *21*(6), 713–736.
Frick, P. J., Lilienfeld, S. O., Ellis, M., Loney, B., & Silverthorn, P. (1999). The association between anxiety and psychopathic traits dimensions in children. *Journal of Abnormal Child Psychology*, *27*(5), 383–392.
Galen, B. R., & Underwood, M. K. (1997). A developmental investigation of social aggression among children. *Developmental Psychology*, *33*, 589–600.
Gustafsson, J. E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, *28*, 407–434.
Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathic traits* (pp. 58–88). New York: Guilford.
Hoge, R. D. (2005). Youth Level of Service/Case Management Inventory. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 283–294). New York: Guilford.
Hoge, R. D., & Andrews, D. A. (2002). *Youth Level of Service/Case Management Inventory User's Manual.* North Tonawanda, NY: Multi-Health Systems.
Kelloway, E. K., & Barling, J. (1990). Item content versus item wording: Disentangling role conflict and role ambiguity. *Journal of Applied Psychology*, *75*(6), 738–742.
Kimonis, E. R., Frick, P. J., Fazekas, H., & Loney, B. R. (2006). Psychopathic traits, aggression, and the processing of emotional stimuli in non-referred children. *Behavioral Sciences and the Law*, *24*, 21–37.
Kimonis, E.R., Marsee, M.A., & Frick, P.J. (2004). The Peer Conflict Scale. Unpublished rating scale.
Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, *64*, 325–347.
Kosson, D. S., Smith, S. S., & Newman, J. P. (1990). Evaluation of the construct validity of psychopathic traits and black and white male inmates: Three preliminary studies. *Journal of Abnormal Psychology*, *99*, 250–259.
Krueger, R. F., Schmutte, P. S., Caspi, A., Moffitt, T. E., Campbell, K., & Silva, P. A. (1994). Personality traits are linked to crime among men and women: Evidence from a birth cohort. *Journal of Abnormal Psychology*, *103*, 328–338.
Kruh, I. P., Frick, P. J., & Clements, C. B. (2005). Historical and personality correlates to the violence patterns of juveniles tried as adults. *Criminal Justice and Behavior*, *32*(1), 69–96.
Lee, Z., Vincent, G. M., Hart, S. D., & Corrado, R. R. (2003). The validity of the Antisocial Process Screening Device as a self-report measure of psychopathy in adolescent offenders. *Behavioral Sciences & the Law*, *21*(6), 771–786.
Little, T. D., Jones, S. M., Henrich, C. C., & Hawley, P. H. (2003). Disentangling the "whys" from the "whats" of aggressive behavior. *International Journal of Behavioral Development*, *27*, 122–133.
Loney, B. R., Frick, P. J., Clements, C. B., Ellis, M. L., & Kerlin, K. (2003). Callous–unemotional traits, impulsivity, and emotional processing in antisocial adolescents. *Journal of Clinical Child and Adolescent Psychology*, *32*, 66–80.
Lorenz, A. R., & Newman, J. P. (2002). Do emotion and information processing deficiencies found in Caucasian psychopaths generalize to African American psychopaths? *Personality and Individual Differences*, *32*, 1077–1086.
Lynam, D. R., Caspi, A., Moffitt, T. E., Raine, A., Loeber, R., & Stouthamer-Loeber, M. (2005). Adolescent psychopathy and the big five: Results from two samples. *Journal of Abnormal Child Psychology*, *33*(4), 431–443.

Munoz, 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.
Pardini, D. A., Lochman, J. E., & Frick, P. J. (2003). Callous/unemotional traits and social–cognitive processes in adjudicated youths. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(3), 364–371.
Patrick, C. J., Hicks, B. M., Nichol, P. E., & Krueger, P. J. (2007). A bifactor approach to modeling the structure of the Psychopathy Checklist—Revised. *Journal of Personality Disorders*, *21*(2), 118–141.
Patrick, C. J., Zempolich, K. A., & Levenston, G. K. (1997). Emotionality and violent behavior in psychopaths: A biosocial analysis. In A. Raine, P. A. Brennan, D. P. Farrington, & S. A. Mednick (Eds.), *Biosocial bases of violence* (pp. 145–161). New York, NY: Plenum Press.
Salekin, R. T., Leistico, A. R., Neumann, C. S., DiCicco, T. M., & Duros, R. L. (2004). Psychopathy and comorbidity in a young offender sample: Taking a closer look at psychopathy's potential importance over disruptive behavior disorders. *Journal of Abnormal Psychology*, *113*(3), 416–427.
Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*(4), 367–373.
Silverthorn, P., Frick, P. J., & Reynolds, R. (2001). Timing of onset and correlates of severe conduct problems in adjudicated girls and boys. *Journal of Psychopathology and Behavioral Assessment*, *23*(3), 171–181.
Skeem, J.L., & Cooke, D.J. (in press). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate.
Skeem, J. L., Mulvey, E. P., & Grisso, T. (2003). Applicability of traditional and revised models of psychopathy to the Psychopathy Checklist: Screening Version. *Psychological Assessment*, *15*(1), 41–55.
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.
Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). *Psychophysiological recording* (2nd ed.). New York: Oxford Press.
Ullman, J. (1996). Structural equation modeling. In B. Tabachnik, & L. Fidell (Eds.), *Using multivariate statistics* (pp. 709–811). New York: Harper Collins.
Vincent, G. M. (2006). Psychopathy and violence risk assessment in youth. *Child and Adolescent Psychiatric Clinics of North America*, *15*(2), 407–428.
Waschbusch, D. A., Pelham, W. E., Jennings, R. J., Greiner, A. R., Tarter, R. E., & Moss, H. B. (2002). Reactive aggression in boys with disruptive behavior disorders: Behavior, physiology, and affect. *Journal of Abnormal Child Psychology*, *30*(6), 641–656.