

The Inventory of Callous-Unemotional Traits: Reliability, Convergent Validity, and Predictive Validity for Reoffending in Adolescents on Probation

Aisha K. Bhanwer Jodi L. Viljoen Catherine S. Shaffer Kevin S. Douglas
Simon Fraser University

Author Note:

Aisha K. Bhanwer, Jodi L. Viljoen, Catherine S. Shaffer, and Kevin S. Douglas,
Department of Psychology, Simon Fraser University.

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada, and a Career Investigator Award to Jodi L. Viljoen from the Michael Smith Foundation for Health Research. The authors would like to thank the youth who participated in this study and the research assistants who assisted with this project.

Correspondence concerning this article should be addressed to Aisha Bhanwer,
Department of Psychology, Simon Fraser University, Burnaby, BC V5A 1S6. Contact:
abhanwer@sfu.ca

Abstract

This prospective study assessed the reliability and validity of the Inventory of Callous-Unemotional Traits (ICU) in a sample of 70 adolescent offenders. With the exception of the Unemotional subscale, ICU scores showed acceptable reliability. ICU total and Callousness scores demonstrated moderate associations with the Antisocial Process Screening Device and the Psychopathy Checklist: Youth Version. The ICU also demonstrated moderate predictive accuracy for violent reoffending over an 18-month follow-up period. While these findings provide some support for the ICU's validity, more research is needed to determine if it is appropriate for use in applied settings with adolescent offenders.

Keywords: adolescent, callous-unemotional, Inventory of Callous-Unemotional Traits, reoffending, violence

The Inventory of Callous-Unemotional Traits: Reliability, Convergent Validity, and Predictive Validity for Reoffending in Adolescents on Probation

Callous-unemotional (CU) traits are characterized primarily by a lack of empathy, shallow affect, and a disregard for the feelings of others (Frick, 2009; Frick, Ray, Thornton, & Kahn, 2014a). CU traits are often viewed as an affective subcomponent of the broader construct of psychopathic traits, which also include behavioral features such as increased impulsivity (Asscher et al., 2011; Cooke & Michie, 2001). Moreover, CU traits in adolescence may be considered an antecedent to psychopathic traits in adulthood (Frick, Wall, Barry, & Bodin, 2016). Adolescents high on CU traits tend to show deficits in emotional processing, such as difficulties with responding to fear or distress (Kimonis, Frick, Fazekas, & Loney, 2006).

They also exhibit greater levels of conduct problems (Frick, 2009) and delinquent behavior (Frick, Ray, Thornton, & Kahn, 2014b). For instance, in a review of 118 studies that examined the association between CU traits and negative outcomes, most studies (89%, $n = 105$) found that CU traits predicted antisocial behavior and aggression (Frick et al., 2014b). Further, 25% ($n = 30$) of studies found that CU traits were related to increased antisocial behavior controlling for conduct problem severity. Studies have also indicated that the presence of greater CU traits is associated with difficulties responding to punishment (Frick et al., 2003).

Given these associations to adverse outcomes, CU traits have recently been included within the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5; American Psychiatric Association, 2013). The DSM-5 includes an addition of a CU specifier entitled *With Limited Prosocial Emotions* to the diagnosis of Conduct Disorder to identify adolescents who exhibit elevated CU traits. An adolescent must exhibit at least two of the four following traits to meet this criterion: lack of remorse or guilt, callousness or lack of empathy, unconcerned about performance, and shallow or deficient affect (American Psychiatric Association, 2013). Further, these characteristics must be pervasive for at least one year.

Despite the clear importance of identifying adolescents with CU traits, it is unclear how clinicians should best screen for these traits. There are several measures designed to assess a broader constellation of psychopathic traits, but few measures focus specifically on CU traits. For instance, one common measure for assessing CU traits is the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003), which was derived from the Psychopathy Checklist – Revised (PCL-R; Hare, 2003). The PCL:YV is a clinician-administered assessment tool that consists of Interpersonal, Affective, Behavioral, and Antisocial factors. In a meta-analytic review, total scores on the PCL:YV were predictive of general ($r_w = .28$) and violent reoffending ($r_w = .25$) over an average follow-up period of 21.9 months (Olver, Stockdale, & Wormith, 2009). Further, psychopathic traits have shown predictive validity across lengthy follow-up periods; a longitudinal study on male adolescent offenders found that PCL:YV scores were predictive of violent reoffending over a 10-year follow-up period (Gretton, Hare, & Catchpole, 2004).

Another tool that is used to assess psychopathic traits, including CU traits, is the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). Like the PCL:YV, the APSD was developed to approximately model the structure and content of the PCL-R (Hare, 2003).

However, in contrast to the PCL:YV, the APSD is a brief measure that, because of its format, requires less time and fewer resources to complete. Although the APSD can be administered to parents and teachers, a self-report version was also developed. The APSD is composed of three factors: Impulsivity, Narcissism, and Callous-Unemotional (CU). In a longitudinal study of 754 at-risk high-school students, self-report APSD CU subscale scores were significantly predictive of both self-reported and official records of offending (McMahon, Witkiewitz, Kotler, & The Conduct Problems Prevention Research Group, 2010). Moreover, this relationship remained significant when controlling for other risk factors, such as diagnoses of Attention Deficit Hyperactivity Disorder (ADHD) and Conduct Disorder.

Although the PCL:YV and APSD have demonstrated predictive validity for reoffending, at least three concerns regarding the use of these tools have been raised. First, the family of PCL measures has been widely criticized for overemphasizing antisocial behavior and providing relatively little emphasis on the core features of psychopathy, such as CU traits (Skeem & Cooke, 2010). For instance, only three items on the PCL:YV assess CU traits (i.e., Lack of Remorse or Guilt, Shallow Affect, Callous/Lack of Empathy; Forth et al., 2003). Given that the APSD is adapted from the PCL measures, these same concerns apply (i.e., only six items on the APSD assess CU traits in adolescents).

Second, with respect to the APSD, some questions have been raised about the validity of self-report measures of CU traits. Some authors have expressed concern that self-report measures of CU traits may be prone to social desirability bias or that an adolescent might underreport CU traits (Breuk, Clauser, Stams, Slot, & Doreleijers, 2007). However, unlike clinician-rated tools, self-report tools are not influenced by raters' bias or subjectivity (Blais, Solodukhin, & Forth, 2014). Furthermore, adolescents might disclose information on a self-report tool that would not otherwise be evident to raters. In fact, despite these concerns about self-report measures, some studies indicate that some self-report measures of CU traits show adequate predictive validity for offending, and concurrent validity with clinician-rated tools (e.g., Feilhauer, Cima, & Arntz, 2012; Fink, Tant, Tremba, & Kiehl, 2012; Kimonis, Kennealy, & Goulter, 2016).

Third, the APSD CU subscale has been found to have poor internal consistency (i.e., Cronbach's alpha [α] less than .70; see Nunnally & Bernstein, 1994) in extant research (e.g., $\alpha = .22$ to $.60$; Ansel, Barry, Gillen, & Herrington, 2015; Goodwin, Sellbom, & Salekin, 2015; Poythress et al., 2006). These findings suggest that the items on the APSD CU subscale may not accurately capture CU traits, and additional items may need to be added to increase reliability (Poythress et al., 2006). Thus, the Inventory of Callous-Unemotional Traits (Frick, 2004) was developed as a more comprehensive measure that is focused on CU traits in adolescents.

The Inventory of Callous-Unemotional Traits

The Inventory of Callous-Unemotional Traits (ICU; Frick, 2004) is a measure of CU traits that was developed as an extension of the CU subscale of the APSD. Specifically, ICU items were developed from four of the six APSD CU subscale items that were related to CU traits in clinical and community adolescent populations (Frick, Bodin, & Barry, 2000). These four items were used to create 24 new items, with half of the items being negatively worded.

Similar to the APSD, teacher, parent, and self-report version of the ICU are available. Factor analyses have suggested both three-factor (Essau, Sasagawa, & Frick, 2006; Kimonis et al., 2008) and five-factor (Feilhauer et al., 2012) models of the self-report ICU. However, the three-factor model of Uncaring, Callousness, and Unemotional subscales has been replicated and typically is used in research (e.g., Byrd, Kahn, & Pardini, 2013; Fanti, Frick, & Georgiou, 2009). Although several studies have investigated the ICU's reliability and validity, findings are limited and mixed, as described below.

Reliability. Internal consistency has been good for the self-report ICU total score in several studies (e.g., Ansel et al., 2015; Byrd et al., 2013; Kimonis, Branch, Hagman, Graham, & Miller, 2013), including studies with forensic samples (Kimonis et al., 2008), but findings have been mixed regarding the internal consistency of subscale scores. Within both community and forensic samples, some research has found acceptable internal consistency for the Callousness (Byrd et al., 2013; Kimonis et al., 2008), Uncaring (Byrd et al., 2013; Kimonis et al., 2008), and Unemotional subscales (Kimonis et al., 2013). In contrast, other studies have found poor internal consistency for the Callousness subscale in community samples (Kimonis et al., 2013) and the Unemotional subscale in forensic samples (Kimonis et al., 2008; Kimonis et al., 2016).

Convergent validity. Thus far, only one study has examined the convergent validity between the self-report ICU and the PCL:YV (Feilhauer et al., 2012). In a study of 70 detained adolescent offenders, ICU total scores were not significantly related to PCL:YV total, Affective, or Interpersonal scores. Limited research has examined the associations between ICU and APSD, and findings have been mixed. One study of incarcerated adolescents found that associations between total scores on the ICU and APSD were significant when the same versions of the measure were compared (i.e., self-report with self-report, parent-report with parent-report), however the ICU and APSD total scores were not significantly correlated when the self-report and parent versions were compared (Fink et al., 2012). In line with these findings, studies with mixed clinical and forensic samples have found that self-report ICU total scores were not significantly related to parent-report APSD total scores (Feilhauer et al., 2012). Overall, the APSD has demonstrated stronger associations with the ICU's Uncaring subscale than the Callous and Unemotional subscales in community (Ansel et al., 2015) and forensic samples (Feilhauer et al., 2012). Additional studies are needed on the convergent validity between these measures to determine if different assessment approaches or tools are capturing similar traits.

Predictive validity. Few studies have examined the ICU's predictive validity for offending and findings also have been mixed. Some studies have been conducted with community samples of young adult males (Byrd et al., 2013; Kahn, Byrd, & Pardini, 2013). In a longitudinal study, Kahn et al. (2013) found that self-report ICU scores in young adults were associated with official records of criminal behavior. Of the individuals who scored high on CU traits, 31.4% offended during the 3.5-year follow-up period, and 12.2% had received more than one criminal charge. Using participants from the same community sample, it was found that self-report ICU scores were related to both self-reported and official records of offending (Byrd et al., 2013). While these findings are promising, they may have limited generalizability. For instance, their samples examined young adults rather than adolescents. Further, the sample was drawn from the community (i.e., school), and concurrent offending, rather than reoffending, was examined. It is beneficial to examine reoffending because tools that measure CU or

psychopathic traits are often used in clinical (forensic) settings to identify adolescents at-risk for reoffending (Viljoen, McLachlan, & Vincent, 2010).

Some studies have examined the ICU's predictive validity in adolescent samples. For instance, in a recent study Ansel et al. (2015) reported positive results for predictive validity of the self-report ICU in a sample of adolescents who had dropped out of school. Scores on the Callousness subscale of the ICU added incremental validity to the prediction of self-reported aggression above and beyond other ICU subscales and existing measures of adolescent psychopathic and CU traits (i.e., the self-report APSD, Youth Psychopathic Traits Inventory, and Childhood Psychopathy Scale). The ICU Callousness scores added unique variance in the prediction of aggression, and accounted for the most variance out of all the subscales in the model. Further, ICU total, Callousness, and Uncaring scores were significantly related to self-reported delinquency and aggression. Some studies have yielded less positive findings regarding the predictive validity of the ICU. For example, a study of incarcerated adolescents found that the self- and parent-report versions of the ICU and APSD and the PCL:YV were not significantly related to the total number of violent or nonviolent convictions (Fink et al., 2012). However, these studies are limited in that these associations were examined concurrently rather than prospectively.

In fact, although a number of studies have examined the ICU, only two prospective studies have examined predictive validity for reoffending in samples of adolescent offenders (Docherty, Boxer, Huesmann, O'Brien, & Bushman, 2017; Kimonis et al., 2016). These studies have focused on incarcerated adolescent rather than adolescents on probation even though probation is, by far, the most common disposition for adolescents who have committed crimes (Alam, 2015). Findings demonstrated that adolescents with higher self-report ICU scores reoffended at a faster rate than those with lower self-report ICU scores over a 2.7-year follow-up period (for both general and violent reoffending; Kimonis et al., 2016). Notably, self-report ICU scores added unique variance in the prediction of reoffending, above and beyond self-reported aggression alone. In another study, ICU total scores on self-, parent-, and teacher-report versions demonstrated large effect sizes in the prediction of offending (Docherty et al., 2017).

In addition to the aforementioned limitations, little research has examined whether ICU scores can be used to discriminate between low- and high-risk adolescents. The ICU manual does not provide cut-off scores that may be used to classify adolescents as high or low on CU traits. However, research that tested potential cut-off scores found that self-report ICU total scores of 28 were able to discriminate their sample of incarcerated adolescents from a school sample (Docherty et al., 2017). While these results provide some support for the use of cut-off scores with the ICU, there is a lack of research that examines differences in ICU scores between adolescent offenders who do and do not reoffend.

Current Study

To address the limitations above, this study used a prospective design to evaluate the reliability and validity of the self-report ICU among a sample of adolescent offenders on probation. First, we examined the internal consistency of ICU total and subscale scores, as existing research has found mixed findings on the subscale scores (e.g., Byrd et al., 2013;

Kimonis et al., 2013). Second, we tested the convergent validity of the ICU with two existing measures that include CU traits, the PCL:YV and APSD. Third, we examined the predictive validity of the ICU, PCL:YV, and APSD for official records of violent and general reoffending, including predictive validity for total scores as well as possible cut-off scores on the ICU. It was hypothesized that ICU total and subscale scores would show modest convergent validity with PCL:YV total and subscale scores, but higher convergent validity with total and subscale scores on another self-report tool, the APSD. We also expected that the strongest associations would be between the Affective and CU subscales of these measures. Finally, it was hypothesized that the ICU total score would show moderate predictive validity for predicting violent and general reoffending.

Methods

Participants. Participants were 70 Canadian adolescent offenders (44 boys and 26 girls) on active probation supervision orders. Sample age ranged between 13 and 18 years ($M = 17.01$, $SD = 1.18$). Of the sample, 40.0% ($n = 28$) were Caucasian/European, 25.7% ($n = 18$) were Aboriginal, 12.9% ($n = 9$) were Asian, 8.6% ($n = 6$) were East Indian, 7.1% ($n = 5$) were Hispanic, 2.9% ($n = 2$) were African, and 2.9% ($n = 2$) were Middle Eastern. The gender and ethnicity compositions of this sample were comparable to other studies with adolescent offenders in Canada (Calverley, Cotter, & Halla, 2010). With respect to index offenses, 62.9% ($n = 44$) of the sample had been convicted for non-sexual violence, 32.9% ($n = 23$) for property offenses, and 2.9% ($n = 2$) for sexual violence. Of the sample, 32.9% ($n = 23$) had prior charges or convictions.

Measures

Self-Report ICU (Frick, 2004). The self-report ICU is a 24-item measure of adolescent CU traits. Responses are given on a four-point Likert scale ranging from 0 (*Not at all true*) to 3 (*Definitely true*). The ICU is composed of three subscales: Callousness (nine items; e.g., “I do not care who I hurt to get what I want”), Unemotional (five items; e.g., “I do not show my emotions to others”) and Uncaring (eight items; e.g., “I feel bad or guilty when I do something wrong”; reverse scored). This three-factor model has been supported in a number of studies (e.g., Kahn et al., 2013). The ICU has been shown to have good reliability and validity, although research is limited. The ICU total score has demonstrated good internal consistency (e.g., $\alpha = .80$; Kahn et al., 2013), but the subscales have shown mixed findings (e.g., Berg et al., 2013; Kimonis et al., 2008; Kimonis et al., 2016; see Results section for internal consistency in present study). Higher scores on the ICU have been associated with self-reported delinquency and aggression (Kimonis et al., 2008), criminal charges (Byrd et al., 2013), and future violent and nonviolent offending (Kahn et al., 2013; Kimonis et al., 2016).

PCL:YV (Forth et al., 2003). The PCL:YV is a clinician-rated tool and is the most commonly used measure of psychopathic traits in adolescents aged 12 to 18 years. Items are rated based upon information from semi-structured interviews and file review. The PCL:YV comprises 20 items in Interpersonal (four items), Affective (four items), Behavioral (five items), and Antisocial (five items) domains. Items are rated on a three-point scale with scores of 0, 1, or 2 (i.e., *item does not apply*, *applies to a certain extent*, or *applies*). Both total scores and factor

scores have shown good reliability in past research (e.g., $\alpha = .89$, $\alpha = .71$ to $.79$; Stockdale, Olver, & Wong, 2010). In the present study, PCL:YV total scores had good internal consistency ($\alpha = .86$), and subscale scores had moderate internal consistency ($\alpha = .67$, $.71$, $.70$, and $.72$, for the Interpersonal, Affective, Behavioral, and Antisocial factors, respectively). Interrater reliability was evaluated using 29 randomly selected cases, which were assessed by having two research assistants present for the interview and file review, but item ratings were completed independently. Interrater reliability of the PCL:YV items was excellent (intraclass correlation coefficient = $.92$, single raters absolute agreement; Shrout & Fleiss, 1979). The PCL:YV has demonstrated good predictive validity for reoffending in meta-analytic studies (e.g., Asscher et al., 2011; Olver et al., 2009), and for various antisocial outcomes, such as institutional violence, instrumental violence, and assault with a weapon (Murrie, Cornell, Kaplan, McConville, & Levy-Elkon, 2004). Further, the PCL:YV has predicted violent and nonviolent reoffending in both boys and girls (Stockdale et al., 2010).

Self-Report APSD (Frick & Hare, 2001). The self-report APSD is a 20-item measure of psychopathic traits in adolescents aged 12 to 18 years. Each item is scored on a three-point scale ranging from 0 (*Not at all true*) to 2 (*Definitely true*). The APSD includes three subscales: Narcissism (seven items, e.g., “I am charming and nice to get what I want”), Impulsivity (five items; e.g., “I do not plan ahead or leave things until the last minute”), and Callous/Unemotional traits (six items; e.g., “My emotions are shallow and fake”). The APSD total score has demonstrated good reliability (e.g., $\alpha = .88$) and validity in a number of studies (e.g., Murrie et al., 2004). However, the CU subscale has demonstrated low internal consistency across multiple studies (Poythress et al., 2006). APSD scores have been related to increased delinquency (Frick, Cornell, Barry, Bodin, & Dane, 2003), violent reoffending, reoffending with weapon-use (Douglas, Epstein, & Poythress, 2008), and increased police contacts (Muñoz & Frick, 2007). Consistent with previous research (Poythress et al., 2006), in the present study, APSD total scores ($\alpha = .83$) and the Narcissism subscale had good internal consistency ($\alpha = .78$), whereas the Impulsivity and CU subscales had low and inadequate internal consistency ($\alpha = .64$, and $.34$, respectively).

Official Records. Outcome data were collected from a provincial corrections database. Reoffense variables included the total number of charges and convictions incurred post-baseline assessment, over a follow up period of 18 months. Both adolescent and adult official records of offending were available. Reoffending data were categorized into violent reoffending and general reoffending. Offenses were considered violent using the following definition: “actual, attempted, or threatened infliction of bodily harm of another person” (Douglas, Hart, Webster, & Belfrage, 2013, pp. 36-37). Nonviolent offenses included breaches, probation violations, property offenses, and drug offenses. Of the sample, 30.3% ($n = 20$) of the participants reoffended, and 16.7% ($n = 11$) had violent reoffenses during the 18-month follow up period.

Procedure

The present study used a prospective design. The ICU was completed by 70 adolescents as part of a larger longitudinal study examining mental health needs, risk, and strengths in adolescents on probation (e.g., Viljoen et al., 2016). Participants were recruited from 11 adolescent probation offices in the greater Vancouver area of British Columbia, Canada.¹

Informed assent was obtained from each adolescent and consent was obtained from his or her parents or guardians prior to data collection. Participants completed the ICU and APSD with trained research assistants (RAs) at their probation office or a quiet location in the community (e.g., coffee shop). In addition, RAs scored the PCL:YV using semi-structured interviews, questionnaire data, and comprehensive file reviews and obtained participants' official records to record reoffending subsequent to the initial assessment. Following completion of data collection, RAs obtained adolescents' official records of reoffending, including charges for violent and general offenses, and the date of their first violent and general reoffense. We were able to obtain records for 66 of the 70 adolescents.

Data Analysis

Reliability. Cronbach's alpha (α) was used to evaluate the internal consistency of ICU total and subscale scores (Field, 2009). Values of $\alpha \geq .70$ represent acceptable reliability, and $\alpha \geq .80$ represent good reliability (Kline, 1999). In addition, mean inter-item correlations (MIC, i.e., the average of the correlations between all items) were calculated to account for the influence of the number of items within a scale (Clark & Watson, 1995). For total scores, acceptable MIC values are between .15 to .20, while subscale score MIC values should fall between .40 and .50 (Clark & Watson, 1995).

Convergent validity. Based on visual examination of the quantile-quantile plots and histograms, ICU, PCL:YV, and APSD total and subscale scores were normally distributed so we used parametric approaches to analyze the data. Pearson bivariate correlations were conducted between total scores on the ICU, PCL:YV, and APSD to assess the convergent validity of the ICU. Because the APSD and PCL:YV contain relatively few items on CU traits, we interpreted correlations that were moderate (i.e., $r = .30$ to $.50$; Cohen, 1988) or large (i.e., $r > .50$; Cohen, 1988) as evidence of convergent validity.

Predictive validity. Receiver Operating Characteristic (ROC) curve analyses were used to determine whether the ICU predicts whether adolescents committed an offense (i.e., violent, general) during the 18-month follow-up period. The area under the curve (AUC) represents the probability that a randomly selected recidivist will have a higher score on the ICU than a randomly selected non-recidivist. AUC values of .556, .639, and .714 represent small, moderate, and large effect sizes, respectively (Rice & Harris, 2005). Hanley and McNeil's (1983) z -test for comparing ROC curves was used to test for significant differences in AUC scores for the ICU, PCL:YV, and APSD total scores. Whereas AUCs examine dichotomous outcome (i.e., whether or not the adolescent was charged with an offense), we also examined whether the ICU predicts speed to offending using Cox proportional hazard regression analyses.

Additional discrimination analyses were conducted to determine sensitivity, the probability that a recidivist had a high ICU total score, and specificity, the probability that a non-recidivist had a low ICU total score (Parikh, Mathai, Parikh, Sekhar, & Thomas, 2008; Singh, 2013). Calibration analyses were conducted to determine the proportion of adolescents with high ICU total scores that reoffended (i.e., positive predictive power [PPP]) and the proportion of adolescents with low ICU total scores that did not reoffend (i.e., negative predictive power [NPP]; Parikh et al., 2008). As the ICU does not have a cut-off score to distinguish between

adolescents scoring low and high on CU traits (Feilhauer et al., 2012; Ray, Frick, Thornton, Steinberg, & Cauffman, 2016), we used cut-off scores at the 25th, 50th, 75th, 90th, and 95th percentiles (i.e., 19.50, 28.00, 32.00, 41.00, and 43.50, respectively). These percentiles have been used in similar analyses with self-report measures that include CU traits (i.e., APSD; Shaffer et al., 2016).

Results

Descriptive Statistics

Descriptive statistics (e.g., means, standard deviations) of ICU, PCL:YV, and APSD total and subscale scores are presented in Table 1. ICU total scores and subscale scores were comparable to other samples of adolescent offenders (e.g., Kimonis et al., 2016). In addition, ICU total scores did not differ significantly between boys ($M = 27.82$, $SD = 10.34$) and girls ($M = 23.94$, $SD = 7.39$), $t(67) = 1.65$, $p = .10$.

--Insert Table 1 about here--

Reliability

ICU total and Uncaring subscale scores had good internal consistency, with Cronbach's α values of .84, and MIC values of .18 and .40, respectively. The Callousness subscale had acceptable reliability ($\alpha = .73$; Kline, 1999). However, the Unemotional scale had poor reliability ($\alpha = .26$). These subscales also demonstrated low internal consistency when the numbers of items within the scale were considered (MIC = .25 and .07, respectively).

Convergent Validity

ICU total scores had significant, small to moderate correlations ($r = .30$ to $.50$; Cohen, 1988) with all PCL:YV scores, excluding the Antisocial factor (see Table 2). As expected, all ICU subscales had small to moderate correlations with the Affective factor of the PCL:YV. Compared to the other ICU subscales, the ICU Callousness subscale had particularly consistent relationships with the PCL:YV factor scores and total scores, with moderate associations with the Affective and Behavioral factors. The ICU Uncaring subscale scores had small (i.e., $r < .30$), significant correlations with the PCL:YV Affective and Behavioral subscales only. Finally, small to moderate significant relationships were observed between the ICU Unemotional subscale scores and almost all PCL:YV scores, excluding the Antisocial factor.

Besides these correlations with the PCL:YV, almost all correlations were significant between ICU scores and APSD scores, with moderate to large correlations (i.e., $r_s > .30$). In general, the ICU demonstrated higher correlations with the APSD than the PCL:YV. The ICU Callousness subscale showed moderate to large correlations with the APSD. Further, moderate to large correlations were observed between the ICU Uncaring scores and all APSD scores.

--Insert Table 2 about here--

Predictive Validity for Reoffending

ROC analyses. ICU total scores significantly predicted violent reoffending over the 18-month follow-up period, with an AUC in the moderate range (AUC = .70; see Table 3). Despite not reaching statistical significance, the ICU total and Unemotional scores demonstrated moderate effect sizes for general reoffending (AUC = .64 and .65, respectively).

In analyses that compared the predictive accuracy between the total scores of the three measures, the PCL:YV total score demonstrated significantly stronger predictive validity than the ICU total score for violent ($z = 2.19, p = .03$) and general reoffending ($z = 2.48, p = .01$). The ICU total score and the APSD total score did not show significant differences in predictive validity for violent ($z = 1.09, p = .28$) or general reoffending ($z = .36, p = .72$). However, the APSD was also significantly outperformed by the PCL:YV total score for predicting general reoffending ($z = 2.45, p = .01$), but not for predicting violent reoffending ($z = 1.64, p = .10$). Of the PCL:YV factors, the Antisocial factor showed the greater effect sizes for both violent and general reoffending (AUCs = .86). Of the APSD subscales, the Impulsivity and CU subscales showed the greatest predictive accuracy (AUCs = .65 to .78).

--Insert Table 3 about here--

Cox proportional hazard regression. In the Cox proportional hazard regression analyses, ICU total scores were a significant predictor of time to the first violent offense (Exp (b) = 1.11, 95% CI [1.01, 1.22], Wald $\chi^2(1) = 4.65, p = .03$). That is, higher ICU total scores were associated with faster violent reoffending. ICU total scores did not significantly predict the time to the first general offense (Exp (b) = 1.06, 95% CI [.97, 1.15], Wald $\chi^2(1) = 1.68, p = .20$).

Sensitivity, specificity, NPP, and PPP. Sensitivity and specificity analyses showed that cut-off scores for the ICU at the 25th and 50th percentiles (i.e., 19.50 and 28.00, respectively) yielded greater sensitivity, and lower values for specificity and positive predictive power. In contrast, higher cut-off scores at the 75th to 95th percentiles (i.e., 32.17 to 43.50) demonstrated greater specificity and positive predictive power, and lower sensitivity. Across the range of cut-off scores, negative predictive power remained high (i.e., 86-91% and 71-77% for violent and general reoffending, respectively). Similar findings were found for both violent and general reoffending (see Table 4).

--Insert Table 4 about here--

Discussion

Although the ICU was developed to address the psychometric limitations of other self-report measures of CU traits, such as the APSD, little research has examined the psychometric properties of the ICU among adolescent offenders. Thus, in the current study, we assessed the reliability, convergent validity, and predictive validity of the ICU in a sample of adolescents on probation.

Reliability

In the current study, ICU total and Uncaring subscale scores demonstrated good internal consistency, and the Callousness subscale showed acceptable internal consistency. In contrast, the Unemotional subscale had poor internal consistency. Other research has found similar findings for the total score, as well as poor internal consistency for the Unemotional subscale (Berg et al., 2013; Kahn et al., 2013; Kimonis et al., 2008; Kimonis et al., 2016; Pechorro, Hawes, Gonçalves, & Ray, 2017). MIC values, which are not sensitive to the number of items in a scale, were also poor for the Unemotional subscale. These findings suggest that the total score may be the most reliable score on the ICU, and that the low reliability of Unemotional subscale is a weakness of the measure. Thus, the ICU may need to be revised in the future, by removing or revising weak items or including additional items to better capture the construct hypothesized to underlie the subscale. Indeed, research with a shortened 12-item version of the parent-report ICU that included only one item from the Unemotional subscale demonstrated good reliability, convergent validity, and discriminant validity in a clinical sample of boys aged 6 to 12 years (Hawes et al., 2014). In addition, a shortened 10-item version of the self-report ICU, with only Callousness and Uncaring subscales, showed comparable reliability and validity to the full version of the ICU in a sample of adolescent offenders (Ray et al., 2016).

The low reliability of the Unemotional subscale might relate to the valence of the items. In particular, the Callousness scale comprises positively valenced items (e.g., “I do not care who I hurt to get what I want”) and the Uncaring subscales comprises negatively-valenced items (e.g., “I feel bad or guilty when I do something wrong”). In contrast, the Unemotional score is composed of both positive and negative items (Ray et al., 2016). This could affect reliability in the sense that the items on the Unemotional subscale may show within-individual variation in responses. Findings indicate that positive items are more likely to be given lower ratings, while negative items tend to have higher ratings (Ray et al., 2016). In addition, the Unemotional scale includes items that reflect a masking of emotions instead of shallow affect (Hawes et al., 2014). In post-hoc analysis in the present study, the Cronbach’s alpha value of the Unemotional subscale increased from .26 to .61 if item 19 (“I am very expressive and emotional”) were dropped; this item appears to refer to the adolescent not wanting to show their emotions.

Convergent Validity

Overall, the ICU appeared to have acceptable convergent validity with existing measures of CU traits and the current results were more promising than those reported in some prior studies (e.g., Feilhauer et al., 2012). Moderate correlations were found between ICU and PCL:YV scores. In particular, the Affective factor of the PCL:YV demonstrated strong correlations with ICU total and subscale scores; however, similar results were found between the ICU total score and the PCL:YV Behavioral factor. The ICU appeared to have stronger associations with the APSD than the PCL:YV, specifically with the APSD CU subscale. These findings are consistent with Fink and colleagues (2012) who found large correlations between the ICU and APSD in a sample of incarcerated adolescents. This finding is likely due to the fact that the ICU and APSD are both self-report ratings, whereas the PCL:YV is rated by assessors.

Predictive Validity for Reoffending

The ICU total score demonstrated moderate predictive validity for violent reoffending. Adolescents with high ICU total scores also committed violent reoffenses at a faster rate than other adolescents. The current findings are consistent with one of the only prospective studies of the ICU's predictive validity for reoffending in incarcerated adolescents (Kimonis et al., 2016). Thus, the ICU may be predictive of violence in not only incarcerated adolescents but also adolescents on probation, who may be lower risk. However, the PCL:YV total score showed significantly stronger predictive validity than the ICU, whereas the APSD appeared to fall between the ICU and the PCL:YV. This suggests that the items on the PCL:YV may provide a more comprehensive and accurate assessment of which adolescents are more likely to reoffend.

Results were less positive for the ICU's predictive accuracy for general reoffending. To be consistent with other studies (Olver et al., 2009), general offending was defined to include all offenses (e.g., violent, property, violations). However, it may be that the moderate AUC score of the ICU total and Unemotional scores predicting general reoffending is being driven by the inclusion of violent offenses within this category.

Research by Kimonis et al. (2016) also demonstrated stronger predictive validity for violence than for general reoffending, as they found that a one standard deviation increase in ICU scores was associated with a 54% increase in violent reoffending, but only a 27% increase in general reoffending. Similarly, findings from Fink et al. (2012) suggested that ICU scores were associated with violent convictions only, and were unrelated to nonviolent convictions in incarcerated adolescents. It may be that CU traits have a greater association with offenses that are perpetrated against others (i.e., violent offenses are more interpersonal than nonviolent offenses). For instance, CU traits may impact interpersonal functioning such that adolescents with these traits are disinhibited from engaging in violence, whereas adolescents without CU traits are inhibited from violence due a concern for others. Indeed, some research has shown that a lack of concern for others has a greater association to engaging in victimization, compared to a lack of empathy (Muñoz, Qualter, & Padgett, 2011).

That said, a number of studies have found that CU traits are predictive of nonviolent offenses. Research with community adolescent samples found small correlations between the ICU and violence and theft (Byrd et al., 2013). Further, psychopathic traits in adolescent offenders have been related to increased probation breaches as well as violent reoffending (Gretton, McBride, Hare, O'Shaughnessy, & Kumka, 2001). Other research has found that ICU total scores add incrementally to the prediction of offenses relating to obstructions of justice (e.g., probation violations), but not for the prediction of violent or property offenses (Kahn et al., 2013).

Although there is a lack of clear guidelines on the potential cut-off scores on the ICU (e.g., Docherty et al., 2017; Ray et al., 2016), the present study revealed that cut-off scores of 19.50 to 28.00 for the ICU total score were better able to classify an adolescent who did reoffend than higher scores. In other words, these cut-offs had high sensitivity. This is consistent with other research that has found that cut-off scores of 28 differentiate forensic samples from community samples (Docherty et al., 2017). The cut-off scores of 19.50 to 28.00 also showed high NPP, which suggests that the ICU may be more useful as a screening tool rather than a stand-alone tool. In contrast, higher cut-off scores from 32.17 to 43.50 were better at classifying

adolescents who did not reoffend (i.e., high specificity). This suggests that modest elevations in ICU total scores may relate to the probability of reoffending during a follow-up period of 18 months.

Implications

These results have several implications for future research and current clinical practice. Given that the ICU total score had good reliability and convergent validity, the ICU may be a reliable and valid measure of CU traits for use in research studies. With the recent addition of the CU specifier to the DSM-5, there is a pressing need for continued research on CU traits, particularly on topics such as predictive validity.

Beyond its use in research settings, the ICU could have value as a screening tool in forensic assessments, such as in violence risk assessments, to screen in youth who require a more comprehensive risk assessment and screen out low-risk youth who not require intensive resources if future research continues to find support. The ICU was created to assess CU traits in adolescents more accurately rather than to assess violence risk per se. However, measures of psychopathic traits, such as the PCL:YV, are often used within adolescent risk assessments. Indeed, 79% of forensic psychologists reported that they assessed psychopathic traits as part of adolescent risk assessments (Viljoen et al., 2010). In the current study, the ICU predicted violent reoffending with moderate to large AUC values. These AUC values are comparable to meta-analytic findings on the predictive validity of leading risk assessment tools, such as the Structured Assessment of Violence Risk in Youth (Borum, Bartel, & Forth, 2006; AUC = .71; Singh, Grann, & Fazel, 2011). That said, other tools that are designed to predict general reoffending and have been validated in research, such as the Youth Level of Service and Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002), may be more appropriate when risk for general reoffending is of interest.

As a screening tool, the ICU may offer some advantages over the PCL:YV, as it takes much less time to administer and is more cost-effective (Kimonis et al., 2016). It also provides a more detailed coverage of CU traits (e.g., 24 vs. 3 items). That said, the PCL:YV contains items that are not included in the ICU (e.g., Behavioral factor), which may explain our findings of stronger predictive accuracy for the PCL:YV (AUC = .83 for PCL:YV total score). It may also be plausible that the broader items included with the PCL:YV better capture risk factors for reoffending (e.g., Behavioral and Antisocial factors). Finally, the PCL:YV allows the clinician to consider behaviors that may not be captured from the ICU or may be underreported due to response bias in clinical settings.

Based on the results of this study, the ICU may also offer some advantages over the APSD, as it has higher reliability for measuring of CU traits (i.e., $\alpha = .84$ for ICU total score measure of CU traits vs. $\alpha = .30$ of CU subscale on APSD; see also Ansel et al., 2015; Fink et al., 2012). That said, the reliability coefficients for the ICU Unemotional subscale were inadequate. However, the ICU has more items that assess CU traits than the APSD, so using the ICU total score may provide a more comprehensive assessment, particularly for the purpose of informing intervention needs.

In addition, administration of the ICU may serve as a reliable and valid measure to help screen for the intensity of treatment needs (Kimonis et al., 2016) and to guide treatment planning (Olver et al., 2009). Research has shown that adolescents with CU traits may have poorer treatment response (e.g., Dadds, Cauchi, Wimalaweera, Hawes, & Brennan, 2012; Haas et al., 2011; O'Neill, Lidz, & Heilbrun, 2003), and may require more comprehensive treatment programs. Further, according to the risk-need-responsivity model (RNR; Bonta & Andrews, 2017), adolescents who pose a greater risk for reoffending should be offered a greater level of treatment services than adolescents who have a lower risk for reoffending. Some research has found that parenting and supervision mediates the association between adolescents with CU traits and their relationships with delinquent peers (Kimonis et al., 2004), so treatment providers may benefit from focusing on family based interventions that improve parenting practices. For instance, Family Functional Therapy (FFT) with high therapist adherence has been shown to reduce violence in adolescents on probation (Sexton & Turner, 2010). In addition, research findings on FFT with adolescent offenders with elevated CU traits demonstrated that violence decreased 6 and 12 months post-treatment (White, Frick, Lawing, & Bauer, 2013). Thus, screening for CU traits with a measure such as the ICU may help determine which adolescents could require extra supports to succeed in treatment settings, and/or benefit from more intensive or specialized forms of treatment.

Strengths and Limitations

The present study has a number of strengths. This study used a prospective design and participants were followed for 18 months using official records. Previous research has typically looked at the relationship between CU traits and offending in a concurrent or retrospective context, but prospective research includes fewer biases and confounding factors, as the assessments are completed before the outcome has occurred.

Despite these strengths, the present study included some limitations. First, the findings may be restricted by low power due a small sample size ($n = 70$). However, many of the associations were significant indicating there was sufficient power to detect convergent and predictive validity. Second, only 26 girls were included in the present study, of which only 26.9% ($n = 7$) reoffended during the follow-up period. Further, only one girl reoffended violently. As such, it was not possible to examine gender differences in reliability and predictive utility. These findings and other research show promise for the use of the ICU with boys (e.g., Kimonis et al., 2016), but more research is needed with girls. Finally, the ICU Unemotional subscale and APSD CU subscale both demonstrated poor reliability. Decreased internal consistency of these subscales may have affected results in convergent and predictive validity analyses. Thus, findings on the validity of the ICU Unemotional and APSD CU subscales should be interpreted with caution.

Conclusion

In sum, the current results provide support for the ICU's reliability and convergent validity with existing measures of CU traits, the PCL:YV and APSD. However, preliminary findings on the ICU's predictive validity for reoffending are mixed. ICU scores were predictive of the presence and speed of violent reoffending over an 18-month follow-up. However, the ICU

did not significantly predict general reoffending, and the PCL:YV demonstrated greater predictive validity overall. More research is needed to determine if self-report tools such as the ICU may be beneficial as a time and cost-efficient method to aid in the prediction and management of violent behavior in adolescent offenders.

Endnotes

¹ With respect to the larger study from which the current sample is drawn, data were collected at baseline, and at 3-, 6-, 9-, 12-month follow ups. The ICU was introduced midway through this larger study and administered at the 6-month follow-up. Of the 508 adolescents informed about the larger study, 163 adolescents did not meet eligibility criteria (i.e., adjudicated for an offense, living in the metropolitan area in which the study was conducted). In addition, 126 adolescents declined to participate, and 26 adolescents could not be reached. Recruitment for data collection occurred over a four-year period from 2008 to 2012.

References

- Alam, S. (2015). Youth court statistics, 2013/2014. *Juristat*. Vol. 35, no. 1. Statistics Canada Catalogue no. 85-002-X ISSN 1209-6393.
- Allen, M. K., & Superle, T. (2016). Youth crime in Canada, 2014. *Juristat*, 36(1), 3-50. Statistics Canada Catalogue no. 85-002-x. Retrieved from <http://www.statcan.gc.ca/pub/85-002-x/2016001/article/14656-eng.pdf>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.
- Ansel, L. L., Barry, C. T., Gillen, C. T. A., & Herrington, L. L. (2015). An analysis of four self-report measures of adolescent callous-unemotional traits: Exploring unique prediction of delinquency, aggression, and conduct problems. *Journal of Psychopathology and Behavioral Assessment*, 37, 207-216. doi:10.1007/s10862-014-9460-z
- Asscher, J. J., van Vugt, E. S., Stams, G. M., Deković, M., Eichelsheim, V. I., & Yousfi, S. (2011). The relationship between juvenile psychopathic traits, delinquency and (violent) recidivism: A meta-analysis. *Journal of Child Psychology and Psychiatry*, 52, 1134-1143. doi:10.1111/j.1469-7610.2011.02412.x
- Berg, J. M., Lilienfeld, S. O., Reddy, S. D., Latzman, R. D., Roose, A., Craighead, L. W., & ... Raison, C. L. (2013). The Inventory of Callous and Unemotional Traits: A construct-validated analysis in an at-risk sample. *Assessment*, 20, 532-544. doi:10.1177/1073191112474338
- Blais, J., Solodukhin, E., & Forth, A. E. (2014). A meta-analysis exploring the relationship between psychopathy and instrumental versus reactive violence. *Criminal Justice and Behavior*, 41(7), 797-821. doi:10.1177/0093854813519629.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). Cincinnati, OH: Anderson.
- Borum, R., Bartel, P. A., & Forth, A. E. (2006). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY): Consultation Edition*. Odessa, FL: Psychological Assessment Resources.
- Breuk, R. D., Clauser, C. A. C., Stams, G. J. J. M., Slot, N. W., & Doreleijers, T. A. H. (2007). The validity of questionnaire self-report of psychopathology and parent-child relationship quality in juvenile delinquents with psychiatric disorders. *Journal of Adolescence*, 30, 761-771. doi:10.1016/j.adolescence.2006.10.003
- Byrd, A. L., Kahn, R. E., & Pardini, D. A. (2013). A validation of the Inventory of Callous-Unemotional Traits in a community sample of young adult males. *Journal of Psychopathology and Behavioral Assessment*, 35, 20-34. doi:10.1007/s10862-012-9315-4
- Calverley, D., Cotter, A., & Halla, E. (2010). Youth custody and community services in Canada, 2008/2009. *Juristat*, 30(1). Statistics Canada Catalogue no. 85-002-x. Retrieved from <http://www.statcan.gc.ca/pub/85-002-x/2010001/article/11147-eng.htm>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. doi:10.1037/1040-3590.7.3.309

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*(2), 171-188. doi:10.1037/1040-3590.13.2.171
- Dadds, M. R., Cauchi, A. J., Wimalaweera, S., Hawes, D. J., & Brennan, J. (2012). Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Research, 199*, 201-207. doi:10.1016/j.psychres.2012.04.033
- Docherty, M., Boxer, P., Huesmann, L. R., O'Brien, M., & Bushman, B. (2017). Assessing callous-unemotional traits in adolescents: Determining cutoff scores for the Inventory of Callous and Unemotional Traits. *Journal of Clinical Psychology, 73*, 257-278. doi:10.1002/jclp.22313
- Douglas, K. S., Epstein, M. E., & Poythress, N. G. (2008). Criminal recidivism among juvenile offenders: Testing the incremental and predictive validity of three measures of psychopathic features. *Law and Human Behavior, 32*(5), 423-438. doi:10.1007/s10979-007-9114-8
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing risk for violence*. Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment, 13*, 454-469. doi:10.1177/1073191106287354
- 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*, 285-298. doi:10.1007/s10862-008-9111-3
- Farrington, D. P., Jolliffe, D., Hawkins, J. D., Catalano, R. F., Hill, K. G., & Kosterman, R. (2003). Comparing delinquency careers in court records and self-reports. *Criminology, 41*, 933-958. doi:10.1111/j.1745-9125.2003.tb01009.x
- 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*, 251-262. doi:10.1016/j.ijlp.2012.04.002
- Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.
- Fink, B. C., Tant, A. S., Tremba, K., & Kiehl, K. A. (2012). Assessment of psychopathic traits in an incarcerated adolescent sample: A methodological comparison. *Journal of Abnormal Child Psychology, 40*, 971-986. doi:10.1007/s10802-012-9614-y
- Forth, A. E., Kosson, D., & Hare, R. D. (2003). *The Hare PCL: Youth Version*. Toronto, ON: Multi-Health Systems.
- Frick, P. J. (2004). *The Inventory of Callous-Unemotional Traits*. Unpublished rating scale. The University of New Orleans.
- Frick, P. J. (2009). Extending the construct of psychopathy to youths: Implications for understanding, diagnosing, and treating antisocial children and adolescents. *The Canadian Journal of Psychiatry, 54*, 803-812. doi:10.1177/070674370905401203

- Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment, 12*, 382-393. doi:10.1037/1040-3590.12.4.382
- 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*, 457-47. doi: 10.1023/A:1023899703866
- Frick, P.J., Cornell, A. H., Bodin, S. D., Dane, H. A., Barry, C. T., & Loney, B. R. (2003). Callous-Unemotional traits and developmental pathways to severe aggressive and antisocial behavior. *Developmental Psychology, 39*, 246- 260. doi: 10.1037/0012-1649.39.2.246
- Frick, P. J., & Hare, R. D. (2001). *Antisocial Process Screening Device: APSD*. Toronto, Canada: Multi-Health Systems.
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014a). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry, 55*, 532-548. doi:10.1111/jcpp.12152
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014b). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140*, 1-57. doi:10.1037/a0033076
- Frick, P. J., Wall, T. D., Barry, C. T., & Bodin, S. D. (2016). Applying the concept of psychopathy to children: Implications for the assessment of antisocial youth. In C. B. Gacono, C. B. Gacono (Eds.) , *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 99-114). New York, NY, US: Routledge/Taylor & Francis Group.
- Goodwin, B. E., Sellbom, M., & Salekin, R. T. (2015). Elucidating the construct validity of the Antisocial Process Screening Device (APSD) in a sample of young adults. *Journal of Psychopathology and Behavioral Assessment, 37*, 1-11. doi:10.1007/s10862-014-9444-z
- Gretton, H. M., Hare, R. D., & Catchpole, R. E. (2004). Psychopathy and offending from adolescence to adulthood: A 10-year follow-up. *Journal of Consulting and Clinical Psychology, 72*, 636-645. doi:10.1037/0022-006X.72.4.636
- Gretton, H. M., McBride, M., Hare, R. D., O'Shaughnessy, R., & Kumka, G. (2001). Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior, 28*(4), 427-449. doi:10.1177/009385480102800403
- Haas, S. M., Waschbusch, D. A., Pelham Jr, W. E., King, S., Andrade, B. F., & Carrey, N. J. (2011). Treatment response in CP/ADHD children with callous/unemotional traits. *Journal of Abnormal Child Psychology, 39*, 541-552. doi:10.1007/s10802-010-9480-4
- Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839-843. doi:10.1148/radiology.148.3.6878708

- Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised*. Toronto, Canada: Multi-HealthSystems.
- Hawes, S. W., Byrd, A. L., Henderson, C. E., Gazda, R. L., Burke, J. D., Loeber, R., & Pardini, D. A. (2014). Refining the parent-reported Inventory of Callous–Unemotional Traits in boys with conduct problems. *Psychological Assessment, 26*(1), 256-266. doi:10.1037/a0034718
- Hoge, R. D., & Andrews, D. A. (2002). *Youth Level of Service/Case Management Inventory: User's manual*. Toronto, Canada: Multi Health Services.
- Kahn, R. E., Byrd, A. L., & Pardini, D. A. (2013). Callous-unemotional traits robustly predict future criminal offending in young men. *Law and Human Behavior, 37*, 87. doi:10.1037/b0000003
- Kimonis, E. R., Branch, J., Hagman, B., Graham, N., & Miller, C. (2013). The psychometric properties of the Inventory of Callous–Unemotional Traits in an undergraduate sample. *Psychological Assessment, 25*, 84-93. doi:10.1037/a0029024
- 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 & the Law, 24*(1), 21-37. doi:10.1002/bsl.668
- Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., & ... Morris, A. S. (2008). Assessing callous-unemotional traits in adolescent offenders: Validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry, 31*, 241-252. doi:10.1016/j.ijlp.2008.04.002
- Kimonis, E. R., Kennealy, P. J., & Goulter, N. (2016). Does the self-report Inventory of Callous-Unemotional Traits predict recidivism? *Psychological Assessment, 28*, 1616-1624, doi:10.1037/pas0000292
- Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.
- Knight, G. P., Little, M., Losoya, S. H., & Mulvey, E. P. (2004). The self-report of offending among serious juvenile offenders: Cross-gender, cross-ethnic/race measurement equivalence. *Youth Violence and Juvenile Justice, 2*, 273-295. doi:10.1177/1541204004265878
- McMahon, R. J., Witkiewitz, K., Kotler, J. S., & The Conduct Problems Prevention Research Group (2010). Predictive validity of callous–unemotional traits measured in early adolescence with respect to multiple antisocial outcomes. *Journal of Abnormal Psychology, 119*, 752-763. doi:10.1037/a0020796
- 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*, 299-312. doi: 10.1111/j.1467-9450.2007.00560.x
- Muñoz, L. C., Qualter, P., & Padgett, G. (2011). Empathy and bullying: Exploring the influence of callous-unemotional traits. *Child Psychiatry and Human Development, 42*, 183-196. doi:10.1007/s10578-010-0206-1
- 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 & the Law, 22*, 49-67. doi:10.1002/bsl.573
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric testing* (3rd ed.). New York: McGraw-Hill.

- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*, 329-353. doi:10.1177/0093854809331457
- O'Neill, M. L., Lidz, V., & Heilbrun, K. (2003). Adolescents with psychopathic characteristics in a substance abusing cohort: Treatment process and outcomes. *Law and Human Behavior*, *27*, 299-313. doi:10.1023/A:1023435924569
- Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology*, *56*, 45-50. doi:10.4103/0301-4738.37595
- Pechorro, P., Hawes, S. W., Gonçalves, R. A., & Ray, J. V. (2017). Psychometric properties of the Inventory of Callous-Unemotional Traits short version (ICU-12) among detained female juvenile offenders and community youths. *Psychology, Crime & Law*, *23*, 221-239. doi:10.1080/1068316X.2016.1239724
- Poythress, N. G., Douglas, K. S., Falkenbach, D., Cruise, K., Lee, Z., Murrie, D. C., & Vitacco, M. (2006). Internal consistency reliability of the self-report Antisocial Process Screening Device. *Assessment*, *13*, 107-113. doi:10.1177/1073191105284279
- Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2016). Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological Assessment*, *28*(4), 394-404. doi:10.1037/pas0000183
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*, 615-620. doi:10.1007/s10979-005-6832-7
- Shaffer, C., Gatner, D., Gray, A. L., Douglas, K. S., Viljoen, J. L., Tweed, R., & ... Gagnon, N. (2016). Incremental and predictive validity of the antisocial process screening device in a community sample of male and female ethnic minority and Caucasian youth. *Journal of Abnormal Child Psychology*, *44*, 1599-1612, doi:10.1007/s10802-016-0130-3
- Sexton, T., & Turner, C. W. (2010). The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. *Journal of Family Psychology*, *24*, 339-348. doi:10.1037/a0019406
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420-428. doi:10.1037/0033-2909.86.2.420
- Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences and the Law*, *31*, 8-22. doi:10.1002/bsl.2052
- Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*, 499-513. doi:10.1016/j.cpr.2010.11.009
- Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, *22*, 433-445. doi:10.1037/a0008512

- Stockdale, K. C., Olver, M. E., & Wong, S. P. (2010). The Psychopathy Checklist: Youth Version and adolescent and adult recidivism: Considerations with respect to gender, ethnicity, and age. *Psychological Assessment, 22*, 768-781. doi:10.1037/a0020044
- Viljoen, J. L., Gray, A. L., Shaffer, C., Bhanwer, A., Tafreshi, D., & Douglas, K. S. (2016). Does reassessment of risk improve predictions? A framework and examination of the SAVRY and YLS/CMI. *Psychological Assessment*. Advance online publication. doi:10.1037/pas0000402
- Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. *Assessment, 17*(3), 377-395. doi:10.1177/1073191109359587
- White, S. F., Frick, P. J., Lawing, K., & Bauer, D. (2013). Callous–unemotional traits and response to functional family therapy in adolescent offenders. *Behavioral Sciences & the Law, 31*, 271-285. doi:10.1002/bsl.2

Table 1
Descriptive Statistics of ICU, PCL:YV, and APSD.

	<i>M (SD)</i>	<i>Range</i>
ICU		
Total	26.42 (9.50)	7.00 - 48.00
Callousness	5.79 (4.08)	0.00 - 17.00
Uncaring	10.41 (5.30)	1.00 - 24.00
Unemotional	7.92 (2.83)	1.00 - 14.00
PCL:YV		
Total	12.83 (6.93)	0.00 - 32.00
Interpersonal	1.50 (1.68)	0.00 - 7.00
Affective	2.23 (1.90)	0.00 - 7.00
Behavioral	3.32 (2.17)	0.00 - 9.00
Antisocial	5.15 (2.68)	0.00 - 10.00
APSD		
Total	14.08 (6.24)	0.00 - 32.00
Narcissism	3.96 (3.00)	0.00 - 12.00
Impulsivity	4.33 (2.05)	0.00 - 9.00
CU	4.35 (1.88)	0.00 - 8.40

Note. ICU = Inventory of Callous-Unemotional Traits. PCL: YV = Psychopathy Checklist: Youth Version. APSD = Antisocial Process Screening Device.

Table 2
Convergent Validity: Pearson Bivariate Correlations Between ICU, PCL:YV, and APSD Scores

	ICU			
	Total	Callousness	Uncaring	Unemotional
PCL:YV				
Total	.43**	.41**	.23	.30*
Interpersonal	.29**	.29**	.14	.27*
Affective	.41**	.36**	.24*	.32*
Behavioral	.39**	.38**	.27*	.29*
Antisocial	.22	.24	.06	.11
APSD				
Total	.69**	.67**	.53**	.35**
Narcissism	.50**	.58**	.33**	.23*
Impulsivity	.42**	.39**	.32*	.24
CU	.77**	.66**	.65**	.50**

Note. ICU = Inventory of Callous-Unemotional Traits. PCL: YV = Psychopathy Checklist: Youth Version. APSD = Antisocial Process Screening Device. * $p < .05$, ** $p < .01$

Table 3
Predictive Validity: ROC Analyses for Presence of Reoffending Over an 18 Month Follow Up

	Violent Reoffending				General Reoffending			
	AUC	SE	<i>p</i>	95% CI	AUC	SE	<i>p</i>	95% CI
ICU								
Total	.70	.10	.042	.51, .88	.64	.08	.068	.49, .80
Callousness	.63	.10	.184	.42, .83	.55	.08	.486	.40, .71
Uncaring	.63	.10	.182	.44, .82	.60	.08	.202	.45, .79
Unemotional	.60	.09	.290	.43, .77	.65	.07	.052	.51, .78
PCL:YV								
Total	.83	.05	.000	.73, .94	.83	.05	.000	.73, .94
Interpersonal	.76	.10	.014	.57, .95	.62	.09	.140	.45, .80
Affective	.83	.07	.001	.69, .97	.69	.08	.022	.53, .85
Behavioral	.83	.08	.001	.67, .99	.79	.07	.000	.66, .92
Antisocial	.86	.06	.000	.75, .98	.86	.05	.000	.76, .95
APSD								
Total	.65	.09	.079	.47, .82	.65	.09	.079	.47, .82
Narcissism	.70	.12	.057	.47, .93	.55	.09	.522	.37, .74
Impulsivity	.78	.10	.007	.59, .97	.65	.08	.081	.49, .81
CU	.74	.08	.021	.59, .89	.69	.08	.023	.54, .84

Note. ICU = Inventory of Callous-Unemotional Traits. PCL: YV = Psychopathy Checklist: Youth Version. APSD = Antisocial Process Screening Device. AUC = area under the curve. SE = standard error. 95% CI = 95% confidence interval of the AUC. *n* = 64. The PCL:YV total score demonstrated significantly stronger predictive validity than the ICU total score for violent ($z = 2.19, p = .03$) and general reoffending ($z = 2.48, p = .01$). The ICU total score and the APSD total score did not show significant differences in predictive validity for violent ($z = 1.09, p = .28$) or general reoffending ($z = .36, p = .72$). The PCL:YV total score significantly outperformed the APSD total score in the prediction of general reoffending ($z = 2.45, p = .01$), but not violent reoffending ($z = 1.64, p = .10$).

Table 4

Table 4
Predictive Accuracy of the ICU Total Score at Various Cut-off Scores

Percentile Cut-Off	Cut-Off Score	Violent Reoffending				General Reoffending			
		Sensitivity	Specificity	PPP	NPP	Sensitivity	Specificity	PPP	NPP
25 th	19.50	.819	.259	.184	.875	.800	.267	.327	.750
50 th	28.00	.727	.574	.258	.912	.600	.578	.387	.765
75 th	32.00	.455	.796	.313	.878	.400	.822	.500	.755
90 th	41.00	.273	.944	.500	.864	.150	.933	.500	.712
95 th	43.50	.182	.981	.667	.855	.100	.978	.667	.710

Note. ICU = Inventory of Callous-Unemotional Traits. PPP = positive predictive power, NPP = negative predictive power.