HHS Public Access

Author manuscript

J Clin Psychol. Author manuscript; available in PMC 2020 May 13.

Published in final edited form as:

J Clin Psychol. 2017 March ; 73(3): 257–278. doi:10.1002/jclp.22313.

Assessing Callous-Unemotional Traits In Adolescents: Determining Cutoff Scores for the Inventory of Callous and Unemotional Traits

Meagan Docherty[1], **Paul Boxer**[2], **L. Rowell Huesmann**[2], **Maureen O'Brien**[2], **Brad Bushman**[3]

[1]Rutgers University-Newark

[2]University of Michigan

[3]Ohio State University

Abstract

Objective: This study aims to establish a cut score for the Inventory of Callous-Unemotional Traits, a well-validated measure of callous-unemotional (CU) traits in youth for which there is currently no cutoff score.

Method: We analyzed data on 634 adolescents from high schools (n = 343) and juvenile detention centers (n = 291). Participants, their parents and guardians, and their teachers and staff members reported on participants' CU traits and aggressive/violent behavior.

Results: All three reports of CU traits as well as intersource composites were associated with aggression, violence, and detained status. Parent report was a better indicator compared to self-reports and teacher reports. Appropriate cut scores based on each report and composite were determined.

Conclusion: We recommend that information from all available informants should be used whenever possible, but when only one informant report is feasible, parent reports are preferable.

Keywords

callous-unemotional traits; aggression; violence; delinquency; adolescents

Introduction

A large number of youth come into contact with the criminal justice system because antisocial and delinquent behavior are more normative and common in the teenage years than during any other developmental phase (Farrington, 1986; Loeber & Farrington, 2014; Moffitt, 1993, 2003; Steffensmeier, Allan, Harer, & Streifel, 1989), although there is some variation based on crime type (Sampson & Laub, 2003; Steffensmeier et al., 1989). For example, in 2011, courts with juvenile jurisdiction handled approximately 1.2 million cases, or roughly 3,400 cases per day (Hockenberry & Puzzanchera, 2014). Whereas most justice-involved youth commit low-level offenses and later desist from offending behavior, a subset

address correspondence to: Dr. Paul Boxer, Department of Psychology, Rutgers University, Newark NJ 07102. pboxer@rutgers.edu.

of youth will exhibit trajectories of problem behavior that are more stable and severe (Moffitt, 1993, 2003); for example, an estimated 26% of the juvenile cases processed in 2011 had committed a person crime (Hockenberry & Puzzanchera, 2014).

It is therefore important for researchers, clinicians, and judicial decision makers to examine factors that distinguish these high-level stable offenders from the rest of their low-level peers. One factor consistently identified in this respect is callous-unemotional (CU) traits, or a constellation of traits that involve a lack of empathy, concern, guilt, remorse, or emotion (Frick, Cornell, Barry, Bodin, & Dane, 2003; Frick, Ray, Thornton, & Kahn, 2014; Frick, Stickle, Dandreaux, Farrell, & Kimonis, 2005; Frick & White, 2008).

Why Are CU Traits Important?

In predicting which antisocial youth will be more persistent and severe in their antisocial behavior, scholars have identified CU traits as one of the key distinguishing factors; youth with CU traits are more likely to commit serious offenses and persist in their offending behavior into adulthood (Frick & Marsee, 2006; Frick et al., 2014; Frick et al., 2005; Frick & White, 2008; Kahn, Byrd, & Pardini, 2013; Lawing, Frick, & Cruise, 2010; McMahon, Witkiewitz, Kotler, & The Conduct Problems Prevention Research Group, 2010). CU traits are often conceptualized or characterized as a “downward” extension of adult psychopathy applied to youth, particularly the affective dimension of psychopathic features (Barry et al., 2000; Frick, 1998; Roose, Bijttebier, Decoene, Claes, & Frick, 2010); therefore, it is not surprising that CU traits in adolescence predict psychopathy in adulthood (Burke, Loeber, & Lahey, 2007).

Indeed, the affective psychopathy dimension (i.e., CU features), more so than the interpersonal, impulsive, or lifestyle dimensions, has consistently predicted which offenders are more severe and persistent (Caputo, Frick, & Brodsky, 1999). Although prevalence rates will differ by population, one multisite study estimated that 2% to 32% of community youth and 14% to 50% of clinic-referred youth meet the criteria for CU traits, depending on whether or not they are diagnosed with conduct disorder and who the informant is (Kahn, Frick, Youngstrom, Findling, & Youngstrom, 2012).

In addition to helping define a subgroup of serious offenders, CU traits are important for understanding which youth might be more resistant to intervention. Previous studies have demonstrated that parenting style is unrelated to conduct problems among youth high in CU traits (Edens, Skopp, & Cahill, 2008; Hawes & Dadds, 2005; Hipwell et al., 2007; Oxford, Cavell, & Hughes, 2003; Wootton, Frick, Shelton, & Silverthorn, 1997), and CU traits are typically associated with poorer treatment outcomes (Frick & Dickens, 2006; Waschbusch, Carrey, Willoughby, King, & Andrade, 2007), more negative behaviors in treatment (Haas et al., 2011), and punishment insensitivity (Barry et al., 2000; Fisher & Blair, 1998; Frick et al., 2003; O’Brien & Frick, 1996). Theorizing on why CU traits may be relatively stable in regards to environmental factors has pointed to the increased heritability of conduct problems among youth with CU traits (Viding, Blair, Moffitt, & Plomin, 2005; Viding, Frick, & Plomin, 2007; Viding, Jones, Frick, Moffitt, & Plomin, 2008).

CU traits also are associated with a number of emotional and cognitive deficits in youth. For example, youth high in CU traits are less concerned about their problem behavior; in fact, they are more likely to expect positive outcomes of their aggression and delinquency (e.g., peer dominance; Pardini & Byrd, 2012; Pardini, Lochman, & Frick, 2003). They also are less accurate in recognizing emotions in facial expressions (Dadds, El Masry, Wimalaweera, & Guastella, 2008; Munoz, 2009; Woodworth & Waschbusch, 2007), are less likely to make eye contact (Dadds et al., 2008) and have decreased amygdala response to distressing stimuli (e.g., fearful faces; Jones, Laurens, Herba, Barker, & Viding, 2009; Marsh et al., 2008). Given that youth with CU traits exhibit greater problem behaviors and attitudes and decreased response to intervention, it is important for researchers, clinicians, and justice officials to be able to determine efficiently which youth are displaying clinically significant levels of these traits.

How Are CU Traits Assessed?

Assessments of CU traits typically involve clinician rating, self-report, parent report, and/or teacher report (Forth, Kosson, & Hare, 2003; Frick, 2003; Frick & Hare, 2001; Kimonis et al., 2008). These assessments can measure the broader construct of youth psychopathy or specifically measure the affective features of psychopathy, or CU traits. Common measures of psychopathy for children and adolescents include the Psychopathy Checklist-Youth Version (PCL-YV; Forth et al., 2003), the Antisocial Process Screening Device (APSD; Frick & Hare, 2001), the Youth Psychopathic Traits Inventory (Andershed, Kerr, Stattin, & Levander, 2002), and the Child Psychopathy Scale (Lynam, 1997).

All of these measures are self-report, except for the PCL-YV, which combines an interview with a file review; the APSD has caregiver and teacher report versions. Although each of these measures is relatively reliable and valid when measuring youth psychopathy (Falkenbach, Poythress, & Heide, 2003; Kosson, Cyterski, Steuerwald, Neumann, & Walker-Matthews, 2002; Munoz & Frick, 2007), there are often only a handful of items within each measure that assess CU traits. For this reason, a measure specifically designed to measure CU traits may be preferred when studying or making decisions based on the affective features of youth psychopathy.

One measure of CU traits specifically is the Inventory of Callous-Unemotional Traits (ICU; Frick, 2003). The ICU was developed based on the original six CU items of the APSD and expands on those items in more detail. It is available in self-report and parent and teacher report versions. Although the entire measure is 24 items, a previous study found 22 of the items to be valid and reliable (Kimonis et al., 2008). The ICU has yielded good internal (Byrd, Kahn, & Pardini, 2013; Essau, Sasagawa, & Frick, 2006; Kimonis et al., 2008), test-retest (Ezpeleta, de la Osa, Granero, Penelo, & Domènech, 2013; Feilhauer, Cima, & Arntz, 2012), and inter-rater reliability (Berg et al., 2013; White, Cruise, & Frick, 2009); it also has been found to have good construct, content, and criterion validity (Byrd et al., 2013; Essau et al., 2006; Kahn et al., 2013; Kimonis et al., 2008). For example, for construct validity, the ICU positively correlates with self-reported aggression and delinquency (Essau et al., 2006; Kimonis et al., 2008) and negatively correlates with empathy and emotional reactivity (Kimonis et al., 2008). For criterion validity, the ICU has demonstrated both concurrent and

predictive validity because it is significantly correlated with both concurrent charges and arrests (Byrd et al., 2013) as well as later charges and arrests (Kahn et al., 2013).

The ICU has been used with a variety of samples, including both detained (Kimonis, Cross, Howard, & Donoghue, 2013; Munoz, Frick, Kimonis, & Aucoin, 2008; Pechorro, Ray, Barroso, Maroco, & Goncalves, 2014) and community samples (Byrd et al., 2013; Ezpeleta et al., 2013; Kimonis, Branch, Hagman, Graham, & Miller, 2013; Roose et al., 2010), as well as in samples with children as young as 3 years (Ezpeleta et al., 2013) to young adults as old as 25 (Byrd et al., 2013). While it does have three subscales (Uncaring, Callousness, and Unemotional) and an acceptable factor structure, the total ICU score consistently has been found to be more reliable than the subscale scores (Kimonis et al., 2008).

Previous studies have not yet examined the criterion validity of the ICU for distinguishing adjudicated and community youth because prior studies principally have administered the ICU to a distinct sample (e.g., adjudicated, clinic-referred) of youth. Therefore, one important step forward in establishing the ICU's criterion validity would be to administer the ICU to two different samples of youth (e.g., adjudicated and community) to determine whether the ICU can reliably discriminate between the two and examine whether a particular cutoff score might be useful in predicting which youth are at greater risk of being detained. A further issue that warrants study is how best to incorporate information from multiple sources; previous studies using the ICU or APSD often use the maximum report across informants (Berg et al., 2013; Bijttebier & Decoene, 2009; Frick, Bodin, & Barry, 2000; Jones et al., 2009; Roose et al., 2010; Viding et al., 2012; White et al., 2009), but further investigation can help determine whether this is the best method across different situations and for various purposes.

The Current Study

CU traits significantly predict aggressive, violent, and delinquent behavior, and they are exhibited by a subgroup of antisocial youth with more severe and stable patterns of problem behavior; further, they can be assessed reliably using the ICU. However, no study has yet examined an appropriate cutoff score for the ICU that would aid in empirical, judicial, or clinical decision making regarding whether youth are exhibiting meaningful levels of CU traits, and there is disagreement about how information from multiple informants should best be used. The current study uses a mixed sample of community and detained adolescents to address two research questions: (a) Can CU traits (as measured by the ICU) reliably distinguish concurrent detained status? And (b) can CU traits predict concurrent aggressive and violent behavior as well as the number of charges and adjudicated offenses and offense seriousness and violence levels among detained youth? We also have two aims in this study: (a) to determine a suitable cutoff score for the ICU based on its criterion validity for predicting whether youth are in the community or detained and (b) how best to use and integrate information from youth, parents, and teachers.

Method

Participants

Data for this study were drawn from the database of a larger project (see Boxer, Huesmann, Bushman, O'Brien, & Moceri, 2009) examining risk factors for violent and nonviolent antisocial behavior among high school students (n = 430) and incarcerated youth (n = 390). In addition to interviews with the youth, information was solicited from parents and guardians as well as teachers and staff. Because the focus of this analysis is on the cross-informant reliability and validity of a rating scale, participants for the current analysis were the 634 youth from the pool of 820 (77.3%), with reports on the target measure (ICU) available from all three sources. In terms of missing data by source, ICU self-reports were missing for four (0.49%) youth, parent and guardian reports were missing for 93 (11.34%) youth, and teacher and staff reports were missing for 106 (12.93%) youth; 17 (2.07%) youth were missing data from two different sources, typically from parent and guardian and teachers and staff (15 youth, 1.83%).

Each informant report on the ICU did not vary as a function of missing data on the other reports; that is, those with missing data on one report did not have significantly higher or lower ICU scores on the other reports. There were no effects of youth sex on the likelihood of reports to be missing from any of the three sources. However, White youth were more likely than were non-White youth to have data available from parents and guardians, $\chi^2(1, 820) = 4.57$, $p < .05$, and from teachers and staff, $\chi^2(1, 820) = 13.05$, $p < .001$. Further, all missing self-reports, $\chi^2(1, 820) = 4.43$, $p < .05$, and more missing parent and guardian reports, $\chi^2(1, 820) = 5.64$, $p < .05$, emanated from the adjudicated sample.

Of the 634 youth retained for analysis (mean [M] age = 16.18 years, standard deviation [SD] = 1.31), 343 (54%) were students and 291 (46%) were detained youth; overall the full sample included 376 males (59%) and 258 females (41%) and a majority of White youth (58% White, 29% Black/African American, 3% Hispanic/Latino/a, 8% multiracial, 2% other). By sample, the student group (M = 16.83, SD = .72) was significantly older ($p < .001$) than was the detained group (M = 15.40, SD = 1.42). Females were significantly under-represented ($p < .001$) in the detained group (28%) relative to the student group (51%); non-Whites were represented about equally across the two groups (39% students, 46% detained youth; $p = .08$).

Measures

CU traits.—Here, we analyze data from the self-rated, parent- and guardian-rated, and teacher- and staff-rated versions of the ICU. Items were the same across all three versions except for the referents of the item stems. All reporters rated "how well" each of 24 statements described the target youth along a 4-point scale ranging from 0 (*not at all true*…) to 3 (*definitely true*). Items tapped three hypothesized components of the CU construct: Uncaring (e.g., "I care about how well I do at school or work"); Callousness (e.g., "I do not care who I hurt to get what I want"); and Unemotional (e.g., "I hide my feelings from others"). The earlier studies of the ICU suggest that scores from the measure may be examined as three separate subfactors of the CU construct as well as global indicators of a

higher order CU factor (i.e., total score). Not surprisingly, these studies observed more meaningful and robust findings with respect to reliability and validity via the total ICU score. Kimonis et al.'s (2008) results also suggested the removal of two consistently unreliable items (does not know right from wrong, does not let feelings control him/her).

Because our focus is on cross-informant reliability and validity rather than factor structure, we used the total ICU scores excising the two unreliable items noted by Kimonis et al. (2008) across self-reports (α = .83), parent and guardian reports (α = .93), and teacher and staff reports (α = .92). In addition to the three separate ICU scores generated by averaging all 22 items for each report, we computed four cross-informant composite scores: (a) the mean composite of all three reports, which was calculated by averaging all three ICU scores (N = 820; Boxer et al., 2009); (b) the mean composite of parent and teacher reports; (c) the max composite of all three reports, which was calculated by taking the highest reported value for each item across all three reports and then summing all items (Frick et al., 2003; Kahn et al., 2012; Piacentini, Cohen, & Cohen, 1992; Roose et al., 2010); and (d) the max composite of parent and teacher reports. The max composite has been used in several previous studies, and although it has been theorized as a better method to aggregate multiple scores from different informers than simply taking the mean, the current study will be the first paper to directly test this idea.

Table 1 displays the percentages of the sample whose max report came from each informant, per ICU item. We also will look at differences between composites of all three reports and composites of parent and teacher reports because youth reports may be more susceptible to social desirability effects (Laajasalo, 2014; Miller & Lynam, 2015), especially given the content of the measure.

Violence and general aggression.—To measure these constructs, we used reliable multi-informant composite scores indicating "violence" (youth (e.g., "How often since you have been a teenager have you punched or beaten someone?"), parent/guardian (e.g., "Using a weapon against another child… How often has this occurred?"), and teacher/staff reports of serious pysical aggression; parent/guardian report of injurious behavior; Boxer et al., 2009; Lefkowitz, Eron, & Walder, 1977) and "General Aggression" not specifically violent in nature (youth report of delinquent behavior (e.g., "How often since you have been a teenager have you… thrown rocks or bottles at people?") and trait aggressiveness (e.g., "If I have to resort to violence to protect my rights, I will"); parent/guardian report of conduct problems (e.g., "Often fights with other youth or bullies them"); teacher/staff report of conduct problems and general aggression (e.g., "What percentage of youth would say that this child… is someone who pushes and shoves others?"); Buss & Perry, 1992; Elliott & Huizinga, 1983; Goodman, 2001; Huesmann, Eron, Guerra, & Crawshaw, 1994).

Study measures used to create these composite variables have been used in adolescent samples similar to ours: (Boxer et al., 2013; Goodman, 2001; Henggeler, Melton, & Smith, 1992; McConville & Cornell, 2003; Morren & Meesters, 2002).

These scores were estimated on the full sample of 820 youth via latent variable modeling using full information maximum likelihood in the AMOS program (version 7.0; Arbuckle,

2006). AMOS applies full information maximum likelihood to analyze the model fit in the presence of missing data and can subsequently generate latent factor scores via regression imputation. Full measurement details including sample items, scale composition, fit statistics, and known-groups validity of factor scores have been described extensively (Boxer et al., 2009). The remaining analyses throughout the paper were conducted in Stata statistical software (StataCorp, 2013).

Offense data.—Data on the offense histories of adjudicated participants were obtained via direct extraction from records held by the partner detention facilities. Trained research assistants copied participants' records by hand, verbatim, from facility files. For each participant, research assistants recorded histories of arrests and associated charges as well as any adjudications and associated charges emanating from arrests. Each charge was coded for seriousness and violence levels following the scheme developed by Rossi, Bose, and Berk (1974; also see Huesmann, Eron, & Dubow, 2002).

Procedures

All procedures were approved by the university's institutional review board (IRB) regulating the implementation of the study, the state agency overseeing the state detention facilities, the federal Office of Human Research Protections, the IRB of the Centers for Disease Control, and the directors or principals of all schools and detention facilities involved in the project. Data collection occurred during 2005 through 2007. Youth were recruited from public high schools (rural, suburban, and urban) and juvenile detention centers (county and state) selected to yield a sample representing a range of risk for aggressive and violent behavior. Across all sites, parent and guardian consent rates averaged about 40% (range by site = 33.6%–48.8%), unsurprising given the length of the survey batteries and nature of the populations sampled. With only slight variations within site types (high school or detention facility), recruitment and interview procedures were conducted differently between the students and detained youth.

In high schools, informational letters and parental consent forms were mailed with stamped return envelopes to parents and guardians of 11th- and 12th-grade students; 2 weeks after the initial mailings, second mailings were sent to parents and guardians who had not responded by that time. Remaining parents and guardians who did not respond by mail to the second contact attempt were solicited by telephone. Parents and guardians could grant permission for their children to participate in writing (mailed) or over the telephone (recorded).

After their children were interviewed, parents, guardians, and youths' teachers (usually social studies teachers) were given survey booklets to complete. Parents and guardians had the option of completing surveys over the telephone, as we have done previously in field research (Boxer et al., 2009). Teachers completed surveys by paper and pencil. Youth interviews were conducted via paper-and-pencil Scantron survey forms in small groups ranging typically from about 10 to 15 students depending upon availability and were led by at least two trained staff for every 10 to 15 students.

In detention facilities, at the start of data collection, the facilities sent informational letters and consent forms typically to parents and guardians of all youth housed, as well as to the

parents and guardians of any new admissions to the facilities over the period of data collection. In these facilities, we were permitted to make the follow-up telephone calls without first sending a second mailing. As with the students, after a detained youth completed his or her interview, we mailed a survey to parents and guardians and provided a survey to staff. Again, parents and guardians had the option to complete their surveys over the telephone, and staff completed surveys by paper and pencil. Trained staff via laptop computers individually conducted youth interviews. Most youth interviews with students and detained youth took approximately one hour.

Across data collection sites, all individuals who provided data were compensated financially in some manner, primarily gift certificates to local merchants, with variations from site to site due to agency regulations or extraneous factors. All high school students received $20, except those in a school collaborating with our research team on another investigation necessitating compensation of $40. All detained youth received $10 compensation due to agency restrictions. Parents and guardians of high school students received $25; parents and guardians of detained youth received $50. All teachers and staff received $5 per completed survey, although this was distributed differently by site due to school or agency policies (e.g., teachers typically received cash but staff had their compensation put into a common fund for staffwide rewards such as appreciation lunches).

Results

Descriptive Statistics and Bivariate Correlations

Table 2 shows descriptive statistics for study variables, separately by sample. We computed a series of two-way analyses of variance (ANOVAs) to examine differences by sex and sample (students vs. detained youth) on ICU scores from youths, parents and guardians, and teachers and staff. These analyses showed generally that detained youth received significantly higher scores than did students, across all informants (all ps < .001; partial η^2 estimates = .07 [youth], .23 [parents and guardians], .12 [teachers and staff]). For youth and teachers and staff, these main effects were qualified somewhat by modest sex by sample interactions; high school students were rated lower than detained youth and females were rated lower than males, with a larger gender difference for students than detained youth (both ps < .05; both partial η^2 estimates = .01). Exploratory t-test analyses indicated no difference in youth and parent and guardian ICU scores as the function of race (White vs. non-White); teachers and staff rated non-White youth higher on the ICU compared to White youth (p < .001). Youth age was modestly to moderately negatively correlated with ICU scores (r values −.17 to −.34, p < .001).

Paired samples t-tests indicated that youth underreported ICU scores compared to parents (p < .001) and teachers (p < .001), but parents and teachers reported relatively similar mean ICU scores (p = .116). Separate paired t tests by sample indicated that informer agreement differed by sample; although youth and parents agreed more for the high school sample (p = .649), youth in the adjudicated sample underreported symptoms compared to their parents (p < .001). Youth underreported symptoms compared to teachers in both samples (p < .001). For the high school sample, teachers reported more symptoms than parents (p < .001), but the reverse was true for the adjudicated sample (p = .024).

Table 3 shows the bivariate correlations among the manifest behavioral criterion measures and the ICU ratings, across informants, separately by sample. Cross-informant correlations on the ICU were modest in magnitude for both samples and slightly smaller for the sample of detained youth. Intra-rater correlations generally are higher than are cross-informant correlations. Still, one typically does not expect great consistency across informants for youth behavior ratings, necessitating the aggregation of cross-informant reports to summarize adequately a target youth's behavioral status. For subsequent analyses, we therefore relied on our cross-informant latent composite scores indicating violence and general aggression.

Using CU Traits to Predict Detained Status

We used a set of logistic regression models to predict detained status, or whether or not youth were detained in a facility, from ICU scores and composites. For each model, age, sex, and race (White/non-White) are included as covariates, and one of the ICU scores (youth, parent, teacher) or composites (mean and max composite of all scores, mean and max composite of parent and teacher scores) is included as a predictor, generating logistic regression models. An eighth model also is included, in which all three ICU scores are entered as predictors. Because non-nested models are being compared, Bayesian information criterion (BIC) and Akaike's information criterion (AIC) scores are reported along with Wald chi-square tests. Cluster-robust standard errors were computed with data collection site as the cluster variable. Because of two cases that were missing data on race/ethnicity, the number of cases drops from 634 to 632 for these models. The results of these logistic regression models are shown in Table 4.

All of the logistic regression models significantly predicted detained status: the model with youth report, Wald $\chi^2(4) = 100.91$, $p < .001$, BIC = 614.91, AIC = 592.67, McFadden's $R^2 = .33$, Tjur's $D = .40$; the model with parent report, Wald $\chi^2(4) = 145.28$, $p < .001$, BIC = 540.05, AIC = 517.81, McFadden's $R^2 = .42$, $D = .49$; the model with teacher report, Wald $\chi^2(4) = 46.56$, $p < .001$, BIC = 605.50, AIC = 583.26, McFadden's $R^2 = .34$, $D = .41$; the model with all three reports as separate predictors, Wald $\chi^2(6) = 444.91$, $p < .001$, BIC = 536.13, AIC = 504.99, McFadden's $R^2 = .44$, $D = .51$; the model with the mean composite of all three reports, Wald $\chi^2(4) = 176.01$, $p < .001$, BIC = 535.99, AIC = 513.75, McFadden's $R^2 = .42$, $D = .50$; the model with the max composite of all three reports, Wald $\chi^2(4) = 180.19$, $p < .001$, BIC = 541.97, AIC = 519.72, McFadden's $R^2 = .41$, $D = .49$; the model with the mean composite of parent and teacher reports, Wald $\chi^2(4) = 88.35$, $p < .001$, BIC = 532.07, AIC = 509.83, McFadden's $R^2 = .43$, $D = .50$; and the model with the max composite of parent and teacher reports, Wald $\chi^2(4) = 82.87$, $p < .001$, BIC = 544.36, AIC = 522.11, McFadden's $R^2 = .41$, $D = .48$.

Because these models are not nested, BIC and AIC can be used to compare the relative fit of these models, with smaller values representing better fit and parsimony for the model. A difference in BIC of less than 2 provides weak evidence that the model with the smaller BIC is a better fit, while a difference of 2 to 6 provides positive evidence, a difference of 6 to 10 is strong evidence, and a difference greater than 10 provides very strong evidence (Raftery, 1995). Using these criteria, the best-fitting model according to BIC appears to be the mean

composite of parent and teacher reports because it has the smallest BIC value by 3.93, which provides positive support. However, AIC indicates that the model with all three reports as separate predictors performs the best, with a difference from the next best model of 4.84, which again provides support for this model as the best-fitting model.

Further, Akaike weights, calculated by taking the difference of each model's AIC and the minimum AIC, indicate that the probability for the model with all three reports as separate predictors has a probability of 90.59, while the other models have a probability of 8.05 or less. The fact that the AIC chose the model with all three reports and the BIC did not is not surprising because this model appeared to perform relatively well compared to the others, and BIC has a greater penalty for model complexity (e.g., more regressors in the model).

To further examine which ICU scores and composites provide a greater advantage in predicting detained status, we conducted receiver operating characteristic (ROC) analyses that plot sensitivity against the inverse specificity (1 = specificity), and we calculated the area under the curve (AUC) for each ROC plot. A significantly greater AUC suggests a measure with a better trade-off between specificity, or the ability to identify positive cases, and sensitivity, or the ability to not identify negative cases. The AUCs are included in Table 4, and Figure 1 shows the ROC curves plotted from each model, except for the covariates-only model (not shown).

The covariates-only model (sex, age, and race) produced an ROC curve (not shown in Figure 1) with an AUC of .85, which was significantly lower than the AUCs of all subsequent models ($p < 0.05$). The remaining models all included these covariates, as well as additional regressors. The next model included the self-report ICU, and produced an ROC curve with an AUC of .86, meaning that a randomly chosen adjudicated participant has an 86% probability of having a higher self-reported ICU score than a randomly chosen high school participant. The remaining AUCs are as follows: for parent report ICU, .90; for teacher report ICU, .86; for all three (youth, parent, and teacher) reports, .90; for the mean composite, .90; and for the max composite, .90.

Significance tests indicated that the self-reports and teacher reports were similar and significantly lower than all other models (besides the covariates-only model; $p < 0.05$), while the parent, max, and mean composites were all similar and significantly higher than the self and teacher reports ($p < 0.05$). The model that included all three reports was significantly higher than all other models ($p < 0.05$), except for the max composite model, from which it was not significantly different. Thus, while BIC values suggest that the model with the mean composite provides the best balance between fit and simplicity, and the AIC values suggest that the model with all three reports provide the most information, the AUC values suggest that the model with the max composite and the model with all three reports both strike the best balance between sensitivity and specificity.

Using CU Traits to Predict Aggressive, Violent, and Antisocial Behavior among Detained Adolescents

To answer our second research question, we used regression models to determine whether CU traits were significantly associated with measures of aggressive and violent behavior,

counts of charges and adjudicated offenses, and seriousness and violence level of crimes, among a sample of detained youth. For each of the four dependent variables, we ran a series of eight regression models with the same set of predictors as in the logistic regressions. The results of these series of models are presented in Table 4 and described below. The number of possible observations for these models drops to 289; out of the 291 detained youth with full ICU information across all informants, two were missing data on race/ethnicity. Again, calculated standard errors are robust in regards to the data collection site. For these models, raw coefficients are reported because standardized coefficients cannot be computed for the count models; however, it should still be easy to compare coefficients across models because all ICU scores and composites are measured on the same scale. The frequencies of charges and adjudicated offenses are reported in Table 5, and the coefficients, significance levels, and BIC values for all models are reported in Table 6.

Because aggressive behavior, violent behavior, and seriousness and violence level of charges were all continuous, normally distributed variables, they are modeled using ordinary least squares regression. For the number of charges and adjudicated offenses, we used negative binomial regression models to account for the overdispersion in the distributions. A zero-truncated negative binomial model is used to predict charges because each of the adjudicated youth necessarily had to have a charge, but this was not the case for adjudicated offenses, for which a regular negative binomial model is used.

In all eight models predicting aggressive behavior ($n = 289$), each of the CU predictors was significant at the $p < .05$ level. According to the BIC values, the three equally best models are the one with all three reports as separate predictors and the mean and max composites of all three reports. The BIC values cannot distinguish among these three because the difference between any two of them is less than two; however, they are preferred to the other models because the next best model has a larger BIC by at least 27.52. According to the Akaike weights, the model with the max composite of all three models is most likely, with a probability of .51, followed by the model with all three reports at .29 and the mean composite model at .20. In all eight models predicting violent behavior ($n = 289$), each of the CU predictors was significant at the $p < .01$ level. According to both BIC and AIC, the best model is the max composite of all three reports. The max composite has the lowest BIC by at least 13.50, and has an Akaike weight probability that rounds to 1.

In the models predicting charges ($n = 245$), none of the individual ICU scores was significant at the $p < .05$ level; however, all four ICU composites were significant at the $p < .001$ level. The BIC values could differentiate only the max composite of all three reports as performing worse than the other composites, while Akaike weights indicated that the two mean composites were equally probable and about three times more probable than either of the max composites. In the models predicting adjudicated offenses ($n = 245$), the only individual ICU score to be significant at the $p < .05$ level was the parent report; however, just as in the models predicting charges, all four composites were significant at the $p < .001$ level.

When considering five models (the model with parent report and the four models with composites as predictors), BIC values indicated that the parent report performed worse than

most composites, but the composites performed relatively similarly, and Akaike weights indicated that the max composite of parent and teacher reports was most probable, with a probability of .31, although the mean composites were close behind at .26 and .25 for all three reports and parent and teacher reports, respectively. When predicting crime seriousness and violence ($n = 245$), none of the CU predictors across all eight models was significant at the $p < .05$ level.

Cutoff Score Analysis to Develop a Dichotomized Index of CU Traits

Finally, two different methods were used to determine the optimal cutoff scores for each measure of ICU, and the results of both are displayed in Table 7. The first method uses predicted probabilities of detained status from each of the models; the predicted probabilities from each model are plotted in Figure 2. Predicted probabilities below .5 indicate a prediction of nondetained status, while predicted probabilities above .5 indicate a prediction of detained status. Therefore, the point at which each measure's marginal effect exceeds .5 can be used as the optimal cutoff score for that measure, above which the likelihood of detained status increases.

Based on these measures, the optimal cutoff scores are as follows: for youth report ICU, 28; for parent report ICU, 30; for teacher report ICU, 33; for the model with all three reports, youth is 40, parent is 30, and teacher is 36; for mean Y/P/T composite, 26.97; for max Y/P/T composite, 42; for mean P/T composite, 28.84; and for max P/T composite, 39. Scores are reported to two decimal places for the mean composites because they are the only scores that have noninteger values. If a more stringent cutoff score is desired to reduce the false positive rate, then the probability cutoff can be increased. Therefore, we also have calculated the cutoff scores for a predicted probability of detained status of .75 as follows: for youth report ICU, 58; for parent report ICU, 48; for teacher report ICU, 61; for the model with all three reports, there is no score available for youth and teacher report because these reports did not produce predicted probabilities beyond .57 and .66, respectively, while the cutoff for parent report ICU is 51; for mean Y/P/T composite, 36.88; for max Y/P/T composite, 56; for mean P/T composite, 40.64; and for max P/T composite, 54.

The other method used to determine appropriate cutoff scores for the ICU was Youden's index, or the sum of each cutoff score's sensitivity and specificity minus one, in which we identified cutoff scores that maximized Youden's index (Youden, 1950). Based on this index, the optimal cutoff scores were as follows: for youth report ICU, 37 (Youden's $J = .60$); for parent report ICU, 41 ($J = .66$); for teacher report ICU, 21 ($J = .61$); for the model with all three scores, the cutoffs are 11 for youth report, 21 for parent report, and 25 for teacher report ($J = .69$); for mean Y/P/T composite, 26.00 ($J = .68$); for max Y/P/T composite, 46 ($J = .66$); for mean P/T composite, 29.50 ($J = .68$); and for max P/T composite, 44 ($J = .65$).

We then classified youth based on these cutoffs and compared AUC values to determine whether there was a difference in using the score from the predicted probabilities or from the Youden's index for each report. Using the cutoff scores generated from the predicted probabilities resulted in more stable and valid results, as indicated by generally greater AUC values, particularly for youth and parent report, the two max composites, and the P/T mean composite. When comparing different scores and composites for the probability cutoffs,

parent report outperformed both youth and teacher report, and both mean composites outperformed both max composites. Parent report and mean composites performed relatively similarly.

Discussion

In this study, we conducted extensive interviews with adolescents in the community as well as detention facilities, obtaining their own self-report data and information from people who knew them well (parents and guardians; teachers and staff members). We used logistic and linear regression models to examine the criterion validity of the ICU among this diverse sample of 634 adolescents to predict detained status as well as its construct validity to predict aggressive and violent behavior, crime seriousness, and violence. As expected, the ICU was significantly associated with concurrent detained status, across all three informants and four different composites, and aggressive and violent behavior among detained youth. However, unexpectedly, ICU scores were not associated with crime seriousness and violence among detained youth.

Further, only composites were associated with charges and adjudicated offenses, except for parent report, which also was associated with adjudicated offenses. Thus, this study makes a significant contribution to the literature by providing this critical evidence of criterion validity for the ICU in its association with detained status and number of charges or offenses and by suggesting that the ICU may not be associated with measures of offending seriousness and violence.

Regarding our first research question in establishing the ICU's validity, we found that CU traits significantly distinguished detained adolescents from high school students. However, we obtained mixed results for our second research question. CU traits were significantly associated with aggressive and violent behavior among detained youth but were not associated with offense seriousness or violence, and for the most part only composites were associated with the number of charges and adjudicated offenses.

Regarding our first aim, we were able to establish cutoff scores for the ICU based on logistic regression models and ROC curves regressing delinquent status on ICU scores and composites. These cutoff scores significantly distinguished between detained adolescents and high school students. Although using such cutoffs would not result in perfect classification, youth with scores above the cutoff are more likely to be similar to detained adolescents and would therefore be more likely to have higher levels of antisocial, aggressive, and violent behavior.

We were able to find evidence to inform our second aim, although again our results depend on the outcome of interest. For example, the max composite (calculating the highest score per item across informants) was preferable when examining violence or the number of adjudicated offenses, but the mean composite (calculating the mathematical average of all reports) was preferable when examining detained status or the number of charges. Regarding individual informant reports, parent report tended to consistently outperform both self and

teacher report; the only exceptions were the number of charges and crime seriousness and violence, in which all three informant reports performed relatively similarly and poorly.

These results are important in identifying where the ICU has criterion and construct validity and where it does not. For example, the ICU may be reliably associated with the number of charges and offenses but not the seriousness or violence level of those offenses. These findings also are important in identifying which informant reports may be more strongly associated with behavioral outcomes because parent report seems to consistently outperform both youth and teacher report. These findings also are essential in establishing cutoff scores for the ICU that can be used for empirical purposes and practical applications. Of course, as with any diagnostic classification system, we are wary of the possible negative effect of labeling youth as being high on a measure of CU traits, and the possible mistake of mischaracterizing a dimensional trait for a taxon. However, the *Diagnostic and Statistical Manual of Mental Disorders Fifth Edition* specifically uses such labeling in its diagnosis of Conduct Disorder (i.e., via the "limited prosocial emotions" specifier; American Psychiatric Association, 2013), and using a dichotomous rather than continuous score can aid clinicians who need to make decisions regarding assessment and intervention.

Although it is perhaps not surprising that the parent report of ICU is a better predictor of detained status than is youth report–because youth might be even less forthcoming to indicate their callous and unemotional personality traits than are their parents–it is somewhat surprising that parent report scores performed better than did teacher report scores. Ostensibly, teachers and staff members should be able to provide the most unbiased view of youths' CU traits because they have nothing to lose or gain from their reporting; they would not feel the same level of shame or social desirability to which youth and parents may be more susceptible. However, this lessened bias comes with a trade-off because teachers and staff members also might be less knowledgeable about youths' true feelings and cognitions. It may be that the report of parents on their children's CU traits may provide the best balance between bias and knowledge: Parents are typically knowledgeable about their children but not as biased against reporting their children's socially undesirable traits as the youths themselves might be.

Limitations

Results from this study offer some insight about the relative contributions of different informants to the assessment of CU traits and how to best integrate this information from multiple sources, but it is not without limitations. First, the data are cross-sectional in nature; therefore, we can provide evidence of only *concurrent,* not *temporal,* predictive validity. It is possible that administering the ICU to youth, parents and guardians, and teachers and staff members *before* the youth were incarcerated would have led to weaker predictions, presumably because the knowledge of a youth as a "delinquent" would shape perceptions of that youth and interpretations of their behavior, leading to inflated ICU scores for the detained youth and an overstated relationship between the two variables.

The cross-sectional design also limits inferences of causality. For example, is it the case that youth high in CU traits are more aggressive and antisocial and therefore end up in detention facilities at higher rates? Or is it instead the case that youth who have been detained in a

facility develop higher CU traits as a response to their environment, compared to youth in the community? Prospective designs following youth over time would lend better insight into this critical issue. Low response rate for the surveys also was an issue and could potentially bias some of the results. Finally, although the ICU captures the four different types of symptoms included in the "with limited prosocial emotions" specifier in the *DSM-5,* we did not attempt to look at each type of symptom specifically to ensure that youth met two or more of the criteria. Future studies would benefit from doing this type of in-depth analysis to aid in approximating the diagnostic criteria, or it could alternatively employ traditional diagnostic assessments and examine their association with "real-world" criteria, such as adjudication status or number or type of offenses.

Conclusion

Despite these limitations, this study has the advantages of a relatively large sample size of 634 youth, a diverse mixed sample of community and detained adolescents, both males and females and from different racial/ethnic backgrounds. Therefore, the results of this study should be more easily generalizable to similar populations. This study also importantly addresses the question of which informant reports may be more useful than others and how to most effectively combine information from multiple informants. This information may be helpful to researchers, clinicians, and judicial decision makers who must assess youths' levels of CU traits, for example, for diagnosis, treatment, investigation, or observation, especially considering the difficulties associated with assessing CU traits (e.g., youth who are callous and unemotional are not necessarily honest in reporting their personality traits).

Our findings might be especially relevant to the ongoing integration of the newest edition of the *DSM* into routine clinical practice in regards to determining whether a youth manifesting Conduct Disorder symptoms also meets criteria for the "limited prosocial emotions" specifier. Clinicians might fruitfully incorporate the ICU into assessments with the youth, the parent or guardian, and a teacher or other adult who might know the youth well. Having all three reports would provide the clinician with a stronger basis for making the diagnosis. However, if budgetary or time constraints prevent the clinician from administering the ICU to all three informants, and only one informant can be approached instead, the clinician should secure an ICU report from the youth's parent or guardian. Once the clinician has obtained the completed ICU from all informants, he or she can determine the likelihood that the youth has CU traits based on the cutoff scores described in this article, and this information can aid in the decision making process of diagnosing the youth. Further research should expand on the temporal predictive validity of the ICU and provide definitive cutoff scores to guide judgments about which youth are experiencing truly clinical or dysfunctional levels of CU traits.

All together, the results of this investigation support the criterion validity of the ICU for differentiating between detained and nondetained youth, but question its construct validity in predicting serious and violent crimes among detained adolescents, while providing relevant information on how to integrate across informants and which scores might be deemed clinically significant.

Acknowledgments

Funding for this study was provided by the U.S. Centers for Disease Control (U49-CE000207; PI: L. R. Huesmann).

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

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

Arbuckle JL (2006). Amos (Version 7.0) [Computer program]. Chicago: SPSS.

Barry CT, Frick PJ, DeShazo TM, McCoy M, Ellis M, & Loney BR (2000). The importance of callous-unemotional traits for extending the concept of psychopathy to children. Journal of Abnormal Psychology, 109(2), 335–340. doi:10.1037/0021-843X.109.2.335 [PubMed: 10895572]

Berg JM, Lilienfeld SO, Reddy SD, Latzman RD, Roose A, Craighead LW,… Raison CL (2013). The Inventory of Callous-Unemotional Traits: A construct-validational analysis in an at-risk sample. Assessment, 20(5), 532–544. doi:10.1177/1073191112474338 [PubMed: 23344913]

Bijttebier P, & Decoene S (2009). Assessment of psychopathic traits in children and adolescents: Further validation of the Antisocial Process Screening Device and the Childhood Psychopathy Scale. European Journal of Psychological Assessment, 25(3), 157–163. doi:10.1027/1015-5759.25.3.157

Boxer P, Huesmann LR, Bushman BJ, O'Brien M, & Moceri D (2009). The role of violent media preference in cumulative developmental risk for violence and general aggression. Journal of Youth and Adolescence, 38, 417–428. doi:10.1007/s10964-008-9335-2 [PubMed: 19636754]

Boxer P, Huesmann LR, Dubow EF, Landau SF, Givirsman SD, Shikaki K, & Ginges J (2013). Exposure to violence across the social ecosystem and the development of aggression: A test of ecological theory in the Israeli-Palestinian conflict. Child Development, 84(1), 163–177. [PubMed: 22906188]

Burke JD, Loeber R, & Lahey BB (2007). Adolescent conduct disorder and interpersonal callousness as predictors of psychopathy in young adults. Journal of Clinical Child & Adolescent Psychology, 36(3), 334–346. doi:10.1080/15374410701444223 [PubMed: 17658978]

Buss AH, & Perry M (1992). The aggression questionnaire. Journal of Personality and Social Psychology, 63, 452–459. doi:10.1037/0022-3514.63.3.452 [PubMed: 1403624]

Byrd AL, Kahn RE, & Pardini DA (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(1), 20–34. doi:10.1007/s10862-012-9315-4

Caputo AA, Frick PJ, & Brodsky SL (1999). Family violence and juvenile sex offending: The potential mediating role of psychopathic traits and negative attitudes toward women. Criminal Justice and Behavior, 26(3), 338–356. doi:10.1177/0093854899026003004

Dadds MR, El Masry Y, Wimalaweera S, & Guastella AJ (2008). Reduced eye gaze explains "fear blindness" in childhood psychopathic traits. Journal of the American Academy of Child & Adolescent Psychiatry, 47(4), 455–463. [PubMed: 18388767]

Edens JF, Skopp NA, & Cahill MA (2008). Psychopathic features moderate the relationship between harsh and inconsistent parental discipline and adolescent antisocial behavior. Journal of Clinical Child & Adolescent Psychology, 37(2), 472–476. doi:10.1080/15374410801955938 [PubMed: 18470783]

Elliott DS, & Huizinga D (1983). Social class and delinquent behavior in a national youth panel: 1976–1980. Criminology, 21(2), 149–177. doi:10.1111/j.174509125.1983.tb00256.x

Essau CA, Sasagawa S, & Frick PJ (2006). Callous-unemotional traits in a community sample of adolescents. Assessment, 3(4), 454–469. doi:10.1177/1073191106287354

Ezpeleta L, de la Osa N, Granero R, Penelo E, & Domenech JM (2013). Inventory of Callous-Unemotional Traits in a community sample of preschoolers. Journal of Clinical Child & Adolescent Psychology, 42(1), 91–105. doi:10.1080/15374416.2012.734221 [PubMed: 23095075]

Falkenbach DM, Poythress NG, & Heide KM (2003). Psychopathic features in a juvenile diversion population: Reliability and predictive validity of two self-report measures. Behavioral Sciences and the Law, 21, 787–805. doi:10.1002/bsl.562 [PubMed: 14696031]

Farrington DP (1986). Age and crime. Crime and Justice, 7, 189–250.

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 [PubMed: 22575180]

Fisher L, & Blair RJR (1998). Cognitive impairment and its relationship to psychopathic tendencies in children with emotional and behavioral difficulties. Journal of Abnormal Child Psychology, 26(6), 511–519. [PubMed: 9915656]

Forth AE, Kosson DS, & Hare RD (2003). The Psychopathy Checklist: Youth Version. Toronto, Canada: Multi-Health Systems.

Frick PJ (1998). Callous-unemotional traits and conduct problems: Applying the two-factor model of psychopathy to children In Cooke DJ, Forth AE & Hare RD (Eds.), Psychopathy: Theory, research and implications for society (pp. 161–187). Netherlands: Kluwer Academic Publishers.

Frick PJ (2003). The Inventory of Callous-Unemotional Traits Unpublished rating scale, The University of New Orleans.

Frick PJ, Bodin SD, & Barry CT (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. doi:10.1037/1040-3590.12.4.382 [PubMed: 11147105]

Frick PJ, Cornell AH, Barry CT, Bodin SD, & Dane HE (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. Journal of Abnormal Child Psychology, 31(4), 457–470. doi:10.1023/A:1023899703866 [PubMed: 12831233]

Frick PJ, & Dickens C (2006). Current perspectives on conduct disorder. Current Psychiatry Reports, 8(1), 59–72. doi:10.1007/s11920-006-0082-3 [PubMed: 16513044]

Frick PJ, & Hare RD (2001). Antisocial process screening device. Toronto, Canada: Multi-Health Systems.

Frick PJ, & Marsee MA (Eds.). (2006). Psychopathy and developmental pathways to antisocial behavior in youth. New York: Guilford.

Frick PJ, Ray JV, Thornton LC, & Kahn RE (2014). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. The Journal of Child Psychology and Psychiatry, 55(6), 532–548. doi:10.1111/jcpp.12152 [PubMed: 24117854]

Frick PJ, Stickle TR, Dandreaux DM, Farrell JM, & Kimonis ER (2005). Callous-unemotional traits in predicting the severity and stability of conduct problems and delinquency. Journal of Abnormal Child Psychology, 33(4), 471–487. doi:10.1007/s10648-005-5728-9 [PubMed: 16118993]

Frick PJ, & White SF (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. Journal of Child Psychology and Psychiatry, 49(4), 359–375. doi:10.1111/j.1469-7610.2007.01862.x [PubMed: 18221345]

Goodman R (2001). Psychometric properties of the strengths and difficulties questionnaire. Journal of the American Academy of Child and Adolescent Psychiatry, 40, 1337–1345. doi:10.1097/00004583-200111000-00015 [PubMed: 11699809]

Haas SM, Waschbusch DA, Pelham WE Jr., King S, Andrade BF, & Carrey NJ (2011). Treatment response in CP/ADHD children with callous/unemotional traits. Journal of Abnormal Child Psychology, 39(4), 541–552. doi:10.1007/s10802-010-9480-4 [PubMed: 21188627]

Hawes DJ, & Dadds MR (2005). The treatment of conduct problems in children with callous-unemotional traits. Journal of Consulting and Clinical Psychology, 73(4), 737–741. doi:10.1037/0022-006X.73.4.737 [PubMed: 16173862]

Henggeler SW, Melton GB, & Smith LA (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. Journal of Consulting and Clinical Psychology, 60(6), 953–961. doi:10.1037/0022-006X.60.6.953 [PubMed: 1460157]

Hipwell AE, Pardini DA, Loeber R, Sembower M, Keenan K, & Stouthamer-Loeber M (2007). Callous-unemotional behaviors in young girls: Shared and unique effects relative to conduct problems. Journal of Clinical Child & Adolescent Psychology, 36(3), 293–304. doi:10.1080/15374410701444165 [PubMed: 17658975]

Hockenberry S, & Puzzanchera C (2014). Juvenile Court Statistics 2011. Pittsburgh, PA: National Center for Juvenile Justice.

Huesmann LR, Eron LD, Guerra NG,& Crawshaw VB (1994). Measuring children's aggression with teachers' predictions of peer-nomination. Psychological Assessment, 6, 329–336. doi:10.1037//1040-3590.6.4.329

Huesmann LR, Eron LD, & Dubow EF (2002). Childhood predictors of adult criminality: Are all risk factors reflected in childhood aggressiveness? Criminal Behaviour and Mental Health, 12(3), 185–208. doi:10.1002/cbm.496 [PubMed: 12830312]

Jones AP, Laurens KR, Herba CM, Barker GJ, & Viding E (2009). Amygdala hypoactivity to fearful faces in boys with conduct problems and callous-unemotional traits. The American Journal of Psychiatry, 166(1), 95–102. [PubMed: 18923070]

Kahn RE, Byrd AL, & Pardini DA (2013). Callous-unemotional traits robustly predict future criminal offending in young men. Law and Human Behavior, 37(2), 87–97. [PubMed: 22731505]

Kahn RE, Frick PJ, Youngstrom E, Findling RL, & Youngstrom JK (2012). The effects of including a callous-unemotional specifier for the diagnosis of conduct disorder. The Journal of Child Psychology and Psychiatry, 53(3), 271–282. doi:10.1111/j.1469-7610.2011.02463.x [PubMed: 21950481]

Kimonis ER, 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(1), 84–93. doi:10.1037/a0029024 [PubMed: 22775408]

Kimonis ER, Cross B, Howard A, & Donoghue K (2013). Maternal care, maltreatment and callous-unemotional traits among urban male juvenile offenders. Journal of Youth and Adolescence, 42(2), 165–177. doi:10.1007/s10964-012-9820-5 [PubMed: 23054349]

Kimonis ER, Frick PJ, Skeem JL, Marsee MA, Cruise KR, Munoz LC,… Morris AS (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 [PubMed: 18514315]

Kosson DS, Cyterski TD, Steuerwald BL, Neumann CS, & Walker-Matthews S (2002). The reliability and validity of the Psychopathy Checklist: Youth Version (PCL:YV) in nonincarcerated adolescent males. Psychological Assessment, 14(1), 97–109. doi:10.1037/1040-3590.14.1.97 [PubMed: 11911052]

Laajasalo T, Aronen ET, Saukkonen S, Salmi V, Aaltonen M, & Kivivuori J (2014). To tell or not to tell? Psychopathic traits and response integrity in youth delinquency surveys. Criminal Behaviour and Mental Health. doi:10.1002/cbm.1940

Lawing K, Frick PJ, & Cruise KR (2010). Differences in offending patterns between adolescent sex offenders high or low in callous-unemotional traits. Psychological Assessment, 22(2), 298–305. doi:10.1037/a0018707 [PubMed: 20528057]

Lefkowitz MM, Eron LD, & Walder LO (1977). Growing up to be violent: A longitudinal study of the development of aggression. New York, NY: Pergamon.

Loeber R,& Farrington DP (2014). Age-crime curve In Bruinsma G &Weisburd D (Eds.), Encyclopedia of Criminology and Criminal Justice (pp. 12–18). New York, NY: Springer.

Lynam DR (1997). Pursuing the psychopath: Capturing the fledgling psychopath in a nomological net. Journal of Abnormal Psychology, 106, 425–438. doi:10.1037//0021-843X.106.3.425 [PubMed: 9241944]

Marsh AA, Finger EC, Mitchell DGV, Reid ME, Sims C, Kosson DS,… Blair RJR (2008). Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional

traits and disruptive behavior disorders. The American Journal of Psychiatry, 165(6), 712–720. doi:10.1176/appi.ajp.2007.07071145 [PubMed: 18281412]

McConville DW, & Cornell DG (2003). Aggressive attitudes predict aggressive behavior in middle school students. Journal of Emotional and Behavioral Disorders, 11(3), 179–187. doi:10.1177/10634266030110030501

McMahon RJ, Witkiewitz K, Kotler JS, &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, 199(4), 752–763. doi:10.1037/a0020796

Miller JD, & Lynam DR (2015). Using self- and informant reports in the assessment of personality pathology in clinical settings: An easy and effective 1–2 combination. Clinical Psychology: Science and Practice, 22(1), 20–24. doi:10.1111/cpsp.12090

Moffitt TE (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. Psychological Review, 100(4), 674–701. doi:10.1037/0033-295X.100.4.674 [PubMed: 8255953]

Moffitt TE (2003). Life-course-persistent and adolescence-limited antisocial behavior: A 10-year research review and a research agenda In Lahey BB, Moffitt TE & Caspi A (Eds.), Causes of conduct disorder and juvenile delinquency (pp. 49–75). New York, NY: Guilford Press.

Morren M, & Meesters C (2002). Validation of the Dutch version of the aggression questionnaire in adolescent male offenders. Aggressive Behavior, 28(2), 87–96. doi:10.1002/ab.90010

Munoz LC (2009). Callous-unemotional traits are related to combined deficits in recognizing afraid faces and body poses. Journal of the American Academy of Child & Adolescent Psychiatry, 48(5), 554–562. doi:10.1097/CHI.0b013e31819c2419 [PubMed: 19318989]

Munoz LC, & Frick PJ (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. [PubMed: 17669220]

Munoz LC, Frick PJ, Kimonis ER, & Aucoin KJ (2008). Types of aggression, responsiveness to provocation, and callous-unemotional traits in detained adolescents. Journal of Abnormal Child Psychology, 36, 15–28. doi:10.1007/s10802-007-9137-0 [PubMed: 17882544]

O'Brien BS, & Frick PJ (1996). Reward dominance: Associations with anxiety, conduct problems, and psychopathy in children. Journal of Abnormal Child Psychology, 24(2), 223–240. [PubMed: 8743246]

Oxford M, Cavell TA, & Hughes JN (2003). Callous/unemotional traits moderate the relation between ineffective parenting and child externalizing problems: A partial replication and extension. Journal of Clinical Child & Adolescent Psychology, 32(4), 577–585. doi:10.1207/S15374424JCCP3204_10 [PubMed: 14710466]

Pardini DA, & Byrd AL (2012). Perceptions of aggressive conflicts and others' distress in children with callous-unemotional traits: 'I'll show you who's boss, even if you suffer and I get in trouble'. Journal of Child Psychology and Psychiatry, 53(3), 283–291. doi:10.1111/j.1469-7610.2011.02487.x [PubMed: 22066467]

Pardini DA, Lochman JE, & Frick PJ (2003). Callous/unemotional traits and social-cognitive processes in adjudicated youths. Journal of the American Academy of Child & Adolescent Psychiatry, 42(3), 364–371. doi:10.1097/01.CHI.0000037027.04952.DF [PubMed: 12595791]

Pechorro P, Ray JV, Barroso R, Maroco J, & Goncalves RA (2014). Validation of the Inventory of Callous-Unemotional Traits among a Portuguese sample of detained juvenile offenders. International Journal of Offender Therapy and Comparative Criminology. doi:10.1177/0306624×14551256

Piacentini JC, Cohen P, & Cohen J (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? Journal of Abnormal Child Psychology, 20, 51–63. doi:10.1007/BF00927116 [PubMed: 1548394]

Raftery AE (1995). Bayesian model selection in social research. Sociological Methodology, 25, 111–163. doi:10.2307/231063

Roose A, Bijttebier P, Decoene S, Claes L, & Frick PJ (2010). Assessing the affective features of psychopathy in adolescence: A further validation of the Inventory of Callous and Unemotional Traits. Assessment, 17(1), 44–57. doi:10.1177/1073191109344153 [PubMed: 19797326]

Sampson RJ, & Laub JH (2003). Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. Criminology, 41(3), 301–340.

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Steffensmeier DJ, Allan EA, Harer MD, & Streifel C (1989). Age and the distribution of crime. American Journal of Sociology, 94(4), 803–831.

Viding E, Blair RJR, Moffitt TE, & Plomin R (2005). Evidence for substantial genetic risk for psychopathy in 7-year-olds. Journal of Child Psychology and Psychiatry, 46(6), 592–597. doi:10.1111/j.1469-7610.2004.00393.x [PubMed: 15877765]

Viding E, Frick PJ, & Plomin R (2007). Aetiology of the relationship between callous-unemotional traits and conduct problems in childhood. British Journal of Psychiatry, 190(49), s33–s38. doi:10.1192/bjp.190.5.s33

Viding E, Jones AP, Frick PJ, Moffitt TE, & Plomin R (2008). Heritability of antisocial behaviour at 9: Do callous-unemotional traits matter? Developmental Science, 11(1), 17–22. doi:10.1111/j.1467-7687.2007.00648.x [PubMed: 18171362]

Viding E, Sebastian CL, Dadds MR, Lockwood PL, Cecil CAM, De Brito SA, & McCrory EJ (2012). Amygdala response to preattentive masked fear in children with conduct problems: The role of callous-unemotional traits. American Journal of Psychiatry, 169(10), 1109–1116. [PubMed: 23032389]

Waschbusch DA, Carrey NJ, Willoughby MT, King S, & Andrade BF (2007). Effects of methylphenidate and behavior modification on the social and academic behavior of children with disruptive behavior disorders: The moderating role of callous/unemotional traits. Journal of Clinsical Child & Adolescent Psychology, 36(4), 629–644. doi:10.1080/15374410701662766

White SF, Cruise KR, & Frick PJ (2009). Differential correlates to self-report and parent-report of callous-unemotional traits in a sample of juvenile sexual offenders. Behavioral Sciences and the Law, 27(6), 910–928. doi:10.1002/bsl.911 [PubMed: 19937922]

Woodworth M, & Waschbusch DA (2007). Emotional processing in children with conduct problems and callous/unemotional traits. Child: Care, Health, and Development, 34(2), 234–244.

Wootton JM, Frick PJ, Shelton KK, & Silverthorn P (1997). Ineffective parenting and childhood conduct problems: The moderating role of callous-unemotional traits. Journal of Consulting and Clinical Psychology, 65(2), 301–308. doi:10.1037/0022-006X.65.2.292.b [PubMed: 9086694]

Youden W J. (1950). Index for rating diagnostic tests. Cancer, 3, 32–35. [PubMed: 15405679]

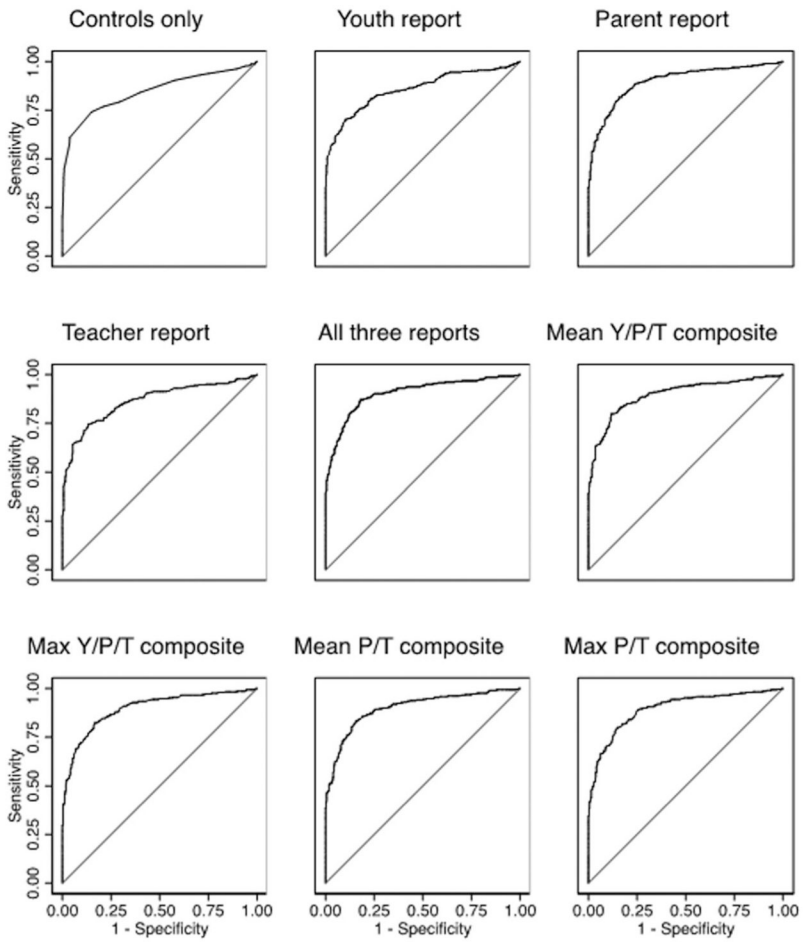

Figure 1.

ROC curves for nine different logistic regression models predicting detained status from controls (age, binary sex, and binary race; included in all models) and various reports and composites of CU traits.

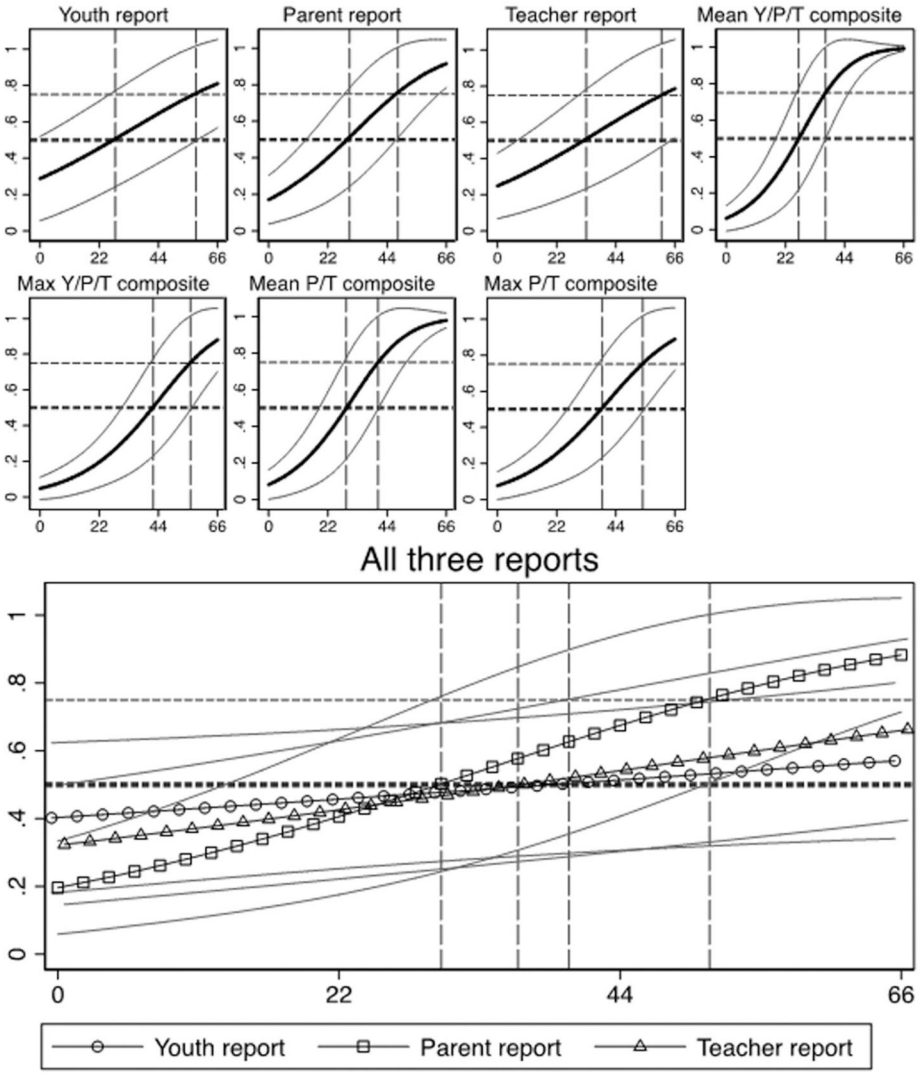

Figure 2.

Predicted probabilities of detained status obtained from logistic regression models, with cutoff scores for ICU reports and composites at which the predicted probability exceeds .5 and .75.

Table 1

Maximum Report Per ICU Item Across Multiple Informants

	Percentage of sample (*n* = 634)						
ICU item	**Youth only**	**Parent only**	**Teacher only**	**Youth and parent**	**Youth and teacher**	**Parent and teacher**	**Youth, parent, and teacher**
1	16.56%	14.83%	18.61%	8.36%	11.83%	12.93%	16.88%
3	4.42%	31.55%	24.92%	5.36%	3.31%	20.98%	9.46%
4	11.20%	20.98%	18.77%	2.68%	2.84%	8.04%	35.49%
5	14.20%	17.98%	24.76%	6.31%	11.67%	14.51%	10.57%
6	33.28%	10.25%	11.83%	10.88%	15.14%	3.00%	15.62%
7	8.36%	22.87%	20.82%	5.68%	5.05%	10.88%	26.34%
8	10.41%	16.88%	26.81%	5.99%	10.57%	16.56%	12.78%
9	16.40%	22.08%	16.88%	5.21%	5.05%	8.52%	25.87%
11	7.26%	23.66%	21.61%	4.26%	4.10%	11.51%	27.60%
12	14.83%	14.67%	17.51%	4.42%	5.84%	6.94%	35.80%
13	14.51%	17.51%	16.56%	9.94%	11.04%	13.88%	16.56%
14	20.98%	12.46%	19.09%	9.94%	13.56%	9.94%	14.04%
15	8.04%	21.92%	26.34%	6.62%	6.94%	18.45%	11.67%
16	10.88%	19.40%	25.55%	7.57%	11.04%	15.14%	10.41%
17	10.25%	22.24%	25.55%	5.36%	8.99%	14.51%	13.09%
18	17.82%	19.40%	17.82%	5.99%	7.26%	7.57%	24.13%
19	17.67%	13.56%	19.09%	9.94%	15.46%	11.04%	13.25%
20	10.25%	27.13%	19.72%	5.21%	4.57%	12.46%	20.66%
21	9.46%	19.87%	23.97%	4.26%	6.15%	9.94%	26.34%
22	25.24%	14.83%	11.04%	9.46%	11.04%	7.26%	21.14%
23	6.94%	25.24%	22.71%	8.99%	5.99%	16.88%	13.25%
24	7.89%	17.82%	23.97%	5.36%	13.56%	17.35%	14.04%

Note. ICU = the Inventory of Callous-Unemotional Traits.

Table 2

Descriptive Statistics for Study Variables by Sample

Variable	N	M	(SD)	Minimum	Maximum
High school students					
Youth ICU	343	19.86	(7.58)	1	44
Parent/guardian ICU	343	20.27	(10.38)	0	66
Teacher/staff ICU	343	23.84	(10.37)	2	54
Mean Y/P/T composite	343	21.32	(6.46)	6.33	40.67
Max Y/P/T composite	343	34.51	(9.18)	12	66
Mean P/T composite	343	22.05	(7.92)	5	44.5
Max P/T composite	343	30.60	(9.86)	10	66
Aggressive behavior	343	−5.89	(5.36)	−14.09	13.83
Violent behavior	343	−.04	(.04)	−.10	.11
Detained youth					
Youth ICU	291	24.88	(9.45)	1	46
Parent/guardian ICU	291	33.24	(12.25)	1	64
Teacher/staff ICU	291	31.26	(9.11)	2	58
Mean Y/P/T composite	291	29.79	(6.58)	14	46.33
Max Y/P/T composite	291	44.96	(8.14)	26	66
Mean P/T composite	291	32.35	(7.82)	10.5	55.5
Max P/T composite	291	41.64	(8.79)	20	64
Aggressive behavior	247	7.03	(8.36)	−13.13	36.31
Violent behavior	247	0.05	(0.06)	−0.10	0.32
Number of charges	247	6.11	(4.46)	1	29
Number adjudicated	247	3.44	(2.93)	0	16
Crime seriousness	247	5.29	(1.07)	3.53	7.25
Crime violence	247	0.53	(0.55)	0	1.73

Note. M = mean; SD = standard deviation; ICU = the Inventory of Callous-Unemotional Traits; Y/P/T = youth, parent, and teacher.

Table 3

Cross-Informant Correlations by Sample, Criterion Measures, and Inventory of Callous-Unemotional Traits (ICU)

Measures/Informant	1	2	3	4	5	6	7	8	9	10	11	12
ICU												
1 Youth	–	.13*	.15*	.41***	.42***	.37*	.09	.01	.02	.18**	.11	.15*
2 Parent and guardian	.19***	–	.05	.05	.02	.07	.73***	.47***	.27***	.05	.09	.00
3 Teacher and staff	.24***	.17**	–	.02	.05	.07	.06	.06	.02	.42***	.25***	.41***
Behavior–Youth												
4 Serious phys. aggression	.28***	.04	.08	–	.79***	.65***	.13*	.05	.11	.24***	.25***	.14*
5 Delinquency	.34***	.10	.12*	.58***	–	.55***	.09	−.08	−.02	.22***	.23***	.15**
6 Trait aggression	.36	.12*	.19***	.57***	.53***	–	.20***	.14*	.15*	.22***	.16**	.20***
Behavior—Parent and guardian												
7 Conduct problems	.12*	.61***	.16**	.06	.14*	.18**	–	.53***	.31***	.19**	.14*	.11
8 Serious phys. aggression behavior	.08	.30***	.11*	.03	.05	.08	.47***	–	.64***	.13*	.07	.09
9 Serious phys. aggression injurious behavior	−.03	.12	.08	.02	.11	.04	.14	.17*	–	.18**	.12	.12
Behavior—Teacher and staff												
10 Conduct problems	.21***	.18**	.63*	.08	.12*	.22***	.22***	.09	.14	–	.57***	.73***
11 Serious phys. aggression	.11*	.10	.28***	.24***	.10	.20***	.09	.04	.07	.55***	–	.54***
12 General aggression	.21***	.11*	.47***	.05	.11	.20***	.08	.03	.04	.73***	.31***	–

Note. High school students (n = 343) below diagonal, detained youth (n = 291) above diagonal.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 4

A Comparison of Logistic Regression Models Predicting Detained Status from ICU Scores and Composites

Predictor	Wald χ^2	BIC	AIC	AUC	Correctly classified	Sensitivity	Specificity	PRE	Cutoff score (50% prob.)	Cutoff score (75% prob.)	Youden's *J* cutoff
Youth ICU	100.91	614.91	592.67	.86	79.59%	73.70%	84.55%	.55	28	58	37
Parent ICU	145.28	540.05	517.81	.90	82.75%	78.55%	86.30%	.62	30	48	41
Teacher ICU	46.56	605.50	583.26	.86	80.85%	74.74%	86.01%	.58	33	61	21
Youth, parent, and teacher ICU	444.91	536.13	504.99	.90	84.18%	79.58%	88.05%	.65	40(y), 30(p), 36(t)	−(y), 51(p), −(t)	11(y), 21(p), 25(t)
Mean Y/P/T composite	176.01	535.99	513.75	.90	84.18%	79.24%	88.34%	.65	26.97	36.88	26.00
Max Y/P/T composite	180.19	541.97	519.72	.90	82.12%	78.55%	85.13%	.61	42	56	46
Mean P/T composite	88.35	532.07	509.83	.90	83.07%	78.89%	86.59%	.63	28.84	40.64	29.50
Max P/T composite	82.87	544.36	522.11	.89	82.59%	78.89%	85.71%	.62	39	54	44

Note. ICU = Inventory of Callous-Unemotional Traits; BIC = Bayesian information criterion; AIC = Akaike information criterion; AUC = area under the curve; PRE = proportion reduction in error; prob. = probability; Y/P/T = youth, parent, and teacher. All Wald χ^2 tests significant at the $p < .001$ level, and all have degree of freedom (*df*) = 4, except for the model with youth, parent, and teacher ICU, which has *df* = 6.

Table 5

Using ICU to Predict Aggressive, Violent, and Antisocial Behavior

	Aggression		Violence		Number of charges		Number adjudicated		Crime seriousness		Crime violence	
Predictor	***b***	**BIC**	***b***	**BIC**	***b***	**BIC**	***b***	**BIC**	***b***	**BIC**	***b***	**BIC**
Youth ICU	.31**	2020.24	.00**	−801.47	.00	1305.05	.01	1114.99	.00	743.14	.00	420.41
Parent ICU	.29**	2011.00	.00**	−818.51	.01	1302.75	.01*	1111.96	.00	742.30	.00	420.00
Teacher ICU	.24**	2039.00	.00**	−784.63	.01	1303.06	.01	1114.59	.00	742.75	.00	420.23
Youth, parent, and teacher ICU	.25***	1960.03	.00**	−859.97	.00	1299.21	.01	1109.31	.00	741.73	.00	419.52
	.24**		.00**		.00		.01		.00		.00	
	.19*		.00**		.01		.01		.00		.00	
Mean Y/P/T composite	.69**	1960.79	.00***	−858.49	.02***	1299.66	.02***	1109.53	.00	742.57	.00	420.22
Max Y/P/T composite	.56***	1958.87	.00**	−873.47	.01***	1302.01	.02***	1110.19	.00	743.18	.00	420.41
Mean P/T composite	.50**	1989.76	.00**	−834.77	.02***	1299.79	.02***	1109.56	.00	741.93	.00	419.76
Max P/T composite	.45**	1988.31	.00**	−841.80	.01***	1301.48	.02***	1109.18	.00	742.49	.00	420.29

Note. ICU = Inventory of Callous-Unemotional Traits; *b* = raw regression coefficients; BIC = Bayesian information criterion; Y/P/T = youth, parent, and teacher; P/T = parent and teacher.Sex, age, and binary race are included as controls in all models. All standard errors were adjusted for clustering by data collection site.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 6

Frequencies of Charges and Adjudicated Offenses

	Charges			Adjudicated offenses		
Count	**Frequency**	**Percentage**	**Cumulative percentage**	**Frequency**	**Percentage**	**Cumulative percentage**
0	0	0.00%	0.00%	24	9.72%	972%
1	24	9.72%	9.72%	46	18.62%	28.34%
2	26	10.53%	20.24%	46	18.62%	46.96%
3	31	12.55%	32.79%	31	12.55%	59.51%
4	23	9.31%	42.11%	34	13.77%	73.28%
5	31	12.55%	54.66%	21	8.50%	81.78%
6	26	10.53%	65.18%	11	4.45%	86.23%
7	15	6.07%	71.26%	11	4.45%	90.69%
8	17	6.88%	78.14%	8	3.24%	93.93%
9	11	4.45%	82.59%	4	1.62%	95.55%
10	10	4.05%	86.64%	3	1.21%	96.76%
11	4	1.62%	88.26%	3	1.21%	97.98%
12	3	1.21%	89.47%	1	0.40%	98.38%
13	7	2.83%	92.31%	1	0.40%	98.79%
14	4	1.62%	93.93%	1	0.40%	99.19%
15	2	0.81%	94.74%	0	0.00%	99.19%
16	5	2.02%	96.76%	2	0.81%	100.00%
17	2	0.81%	97.57%			
18	1	0.40%	97.98%			
19	3	1.21%	99.19%			
22	1	0.40%	99.60%			
29	1	0.40%	100.00%			

Table 7

Using ICU Cutoff Scores to Predict Detained Status

Report	Score	Sensitivity	Specificity	Correctly classified	OR	AUC
Predicted probabilities (>.5)						
Youth ICU	28	40.89%	84.26%	64.35%	3.70	.63***
Parent ICU	30	65.64%	83.97%	75.55%	10.00	.75***
Teacher ICU	33	48.80%	79.59%	65.46%	3.72	.64
Mean composite Y/P/T ICU	26.97	67.35%	78.13%	73.19%	7.37	.73
Max composite Y/P/T ICU	42	62.20%	76.97%	70.19%	5.50	.69**
Mean composite P/T ICU	28.84	68.73%	79.30%	74.45%	8.42	.74*
Max composite P/T ICU	39	60.14%	79.30%	70.50%	5.78	.70***
Youden's index						
Youth ICU	37	10.65%	98.25%	58.04%	6.70	.54
Parent ICU	41	24.74%	95.04%	62.78%	6.30	.60
Teacher ICU	21	89.35%	39.94%	62.62%	5.58	.65
Mean composite Y/P/T ICU	26.00	70.10%	72.89%	71.61%	6.30	.71
Max composite Y/P/T ICU	46	42.27%	87.46%	66.72%	5.12	.65
Mean composite P/T ICU	29.50	65.29%	79.88%	73.19%	7.47	.73
Max composite P/T ICU	44	39.17%	90.09%	66.72%	5.85	.65

Note. OR = odds ratio; AUC = area under the curve; ICU = Inventory of Callous-Unemotional Traits Y/P/T = youth, parent, and teacher. Robust, clustered standard errors were used in computing significance of odds ratios. Significance stars are used for comparisons between AUC for the two types of cutoff scores, with stars denoting the cutoff score with the greater AUC.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.