Short-Term Stability of Psychopathic Traits in Adolescent Offenders

Zina Lee

Department of Psychology, The University of Alabama

Jessica R. Klaver

Kirby Forensic Psychiatric Center, New York, NY

Stephen D. Hart, Marlene M. Moretti, and Kevin S. Douglas

Department of Psychology, Simon Fraser University

There is considerable debate about the assessment of psychopathic traits in adolescence due in part to questions regarding the stability of traits. We investigated the 6-month stability of psychopathic traits in a sample of 83 male adolescent offenders using an augmented protocol for the Psychopathy Checklist: Youth Version and the self-report Antisocial Process Screening Device. Findings suggested moderate to high stability of psychopathic traits, as indexed by total scores, and low to moderate stability of psychopathic traits at the factor level. The interpersonal and behavioral traits demonstrated greater stability relative to the affective traits, and stability varied by developmental stage, with lower stability in early adolescence. Implications for understanding the developmental expression of psychopathic traits in adolescence, as well as for clinical-forensic practice, are discussed.

Extensive research using the Hare Psychopathy Checklist–Revised (PCL–R; Hare, 1991, 2003) illustrates the importance of psychopathy for managing individuals in

forensic and criminal populations. For example, psychopathic offenders commit more general and violent crime, and are less motivated and less responsive to treatment (Douglas, Vincent, & Edens, 2006; Harris & Rice, 2006). As such, interest has extended downward into investigating child and adolescent psychopathic traits to identify youth at high risk for serious and violent delinquency and potentially address the etiology of the disorder.

Much of the research assessing adolescent psychopathic traits has used the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003), and there is substantial evidence for reliability and validity. Psychopathic traits in adolescents demonstrate moderate to large associations with externalizing disorders, such as conduct disorder and attention deficit hyperactivity disorder, but smaller associations with internalizing disorders, such as anxiety and depression (Salekin, Neumann, Leistico, DiCicco, & Duros,

2004). Furthermore, adolescent psychopathic traits are associated with violence, recidivism, and treatment problems (e.g., Gretton, Hare, & Catchpole, 2004; Murrie, Cornell, Kaplan, McConville, & Levy-Elkon, 2004; O'Neill, Lidz, & Heilbrun, 2003). However, this should be viewed in light of recent evidence of considerable heterogeneity in the strength of the association between adolescent psychopathic traits and recidivism (e.g., Edens, Campbell, & Weir, 2007), and evidence that interventions do benefit psychopathic adolescents (e.g., Caldwell, Skeem, Salekin, & van Rybroek, 2006).

Despite evidence for the validity of adolescent psychopathic traits, there are lingering concerns about the appropriateness of assessing such traits in adolescence. Some suggest it may be difficult to identify stable personality traits in childhood or adolescence (Hart, Watt, & Vincent, 2002). It is important to note that researchers have highlighted that abilities, such as awareness of long-term consequences, perspective taking, and emotion regulation, are still in a state of considerable flux and development during adolescence (Edens, Skeem, Cruise, & Cauffman, 2001; Steinberg & Cauffman, 1996; Westen & Chang, 2000), which may influence the development of personality traits. In sum, considerable debate about adolescent psychopathic traits highlights both the potential benefits (e.g., prevention) and consequences (e.g., labeling) with respect to assessment, and the need to further our understanding of the construct by incorporating principles from personality and developmental psychopathology (Frick, 2002; Lynam, 2002; Seagrave & Grisso, 2002).

STABILITY OF PSYCHOPATHIC TRAITS IN ADULTS

With respect to adult psychopathy, there is evidence that stability varies across different time intervals. Schroeder, Schroeder, and Hare (1983) reported a 10-month stability estimate of .89 for PCL–R total scores in a sample of incarcerated male offenders. In samples of men undergoing treatment for substance use, 1-month stability estimates (Pearson rs) ranged from .76 to .89 for the PCL–R total and factor scores (Alterman, Cacciola, & Rutherford, 1993), and 2-year stability estimates (intraclass correlation coefficients [ICCs]) ranged from .43 to .60 for PCL–R total scores (Rutherford, Cacciola, Alterman, McKay, & Cook, 1999).

Although these studies suggest psychopathy is at least moderately stable in adults, this may reflect an artifact of the assessment procedure. The PCL–R is an observer rating scale, and most item ratings reflect lifetime functioning according to an interview and a review of case history information. The lifetime time frame makes it relatively insensitive to fluctuations in symptom severity over time, as any follow-up assessments are based largely on the same case history information on which initial assessments were based. The same concern holds for the PCL:YV (Frick, Kimonis, Dandreaux, & Farell, 2003), as a downward extension of the PCL–R.

STABILITY OF PSYCHOPATHIC TRAITS IN ADOLESCENTS

Personality disorder traits, including those of psychopathy, do not appear suddenly or arise de novo in adulthood. It is generally recognized that they emerge during the course of childhood and adolescence, and may constitute a full-fledged clinical syndrome in adulthood (American Psychiatric Association, 2000 [*Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.)]; World Health Organization, 1992 [International Statistical Classification of Diseases and Related Health Problems–10th Revision]). Research on the stability of personality and personality disorder traits in adolescents demonstrate variability. Broad personality traits, such as extraversion and neuroticism, demonstrate stability correlations ranging from .43 to .65 across childhood and adulthood (e.g., Bazana & Stelmack, 2004; Roberts & DelVecchio, 2000). In contrast, personality disorders demonstrate 2-year stability estimates of .20 for diagnoses (Mattanah, Becker, Levy, Edell, & McGlashan, 1995) and range from .37 to .57 for borderline, histrionic, and narcissistic traits (Johnson et al., 2000). However, there is also evidence that the stability of personality traits varies across development, with lower stability in childhood and adolescence, ranging from .43 to .46, and higher stability in adulthood, ranging from .54 to .65 (e.g., Bazana & Stelmack, 2004; Roberts & DelVecchio, 2000). It is quite likely that the same lower stability in childhood and adolescence would hold true for psychopathic traits.

Researchers have recently begun to examine the stability of child and adolescent psychopathic traits using teacher and parent ratings. Studies of children indicate stability ranges from .59 to .89 across periods of 1 to 4 years (Barry, Barry, Deming, & Lochman, 2008; Frick et al., 2003; Hawes & Dadds, 2007; Pardini, Lochman, & Powell, 2007). In one of the first studies, Frick et al. (2003) examined the 2- to 4-year stability of the parent-rated Antisocial Process Screening Device in a community sample of children. Stability ranged from .80 to .88 for total scores and was only slightly lower for specific trait clusters: .77 to .88 for narcissism, .71 to .86 for callous-unemotional traits, and .72 to .86 for impulsivity. However, there was also evidence of instability. For example, only 30% of those rated high on callous-unemotional traits were rated high across all three follow-up assessments.

Studies examining developmental periods of childhood into adolescence and across adolescence have also found that stability varies. Obradovic, Pardini, Long, and Loeber (2007) found stability estimates ranging from .27 to .84 and evidence of metric variance (i.e., the items functioned differently for children and adolescents), suggesting potential developmental differences. Similarly, Pardini and Loeber (2008) found that although trajectories of interpersonal callousness were generally flat, there was a decrease across time and variability in individual trajectories. Finally, several studies have examined stability from childhood and adolescence into adulthood. Burke, Loeber, and Lahey (2007) found that teacher ratings of interpersonal callousness in childhood predicted PCL–R Factors 1 and 2 in young adulthood. Loney, Taylor, Butler, and Iacono (2007) found 6-year stability estimates ranging from .40 to .41 across late adolescence and early adulthood using self-report measures, whereas Blonigen, Hicks, Krueger, Patrick, and Iacono (2006) found 7-year test–retest correlations of .60 for the interpersonal-affective traits and .53 for the social deviance traits. Lynam, Caspi, Moffitt, Loeber, and Stouthamer-Loeber (2007) found a stability estimate of .31 for psychopathic traits across 11 years from adolescence into adulthood.

THE CURRENT STUDY

The growing body of literature suggests that child and adolescent psychopathic traits exhibit at least moderate stability. Of importance, there is also evidence that the traits are not immutable. However, several outstanding issues remain. First, no studies have examined forensic samples, limiting our understanding of stability in this important group of youth. Studies investigating children and adolescents in forensic populations will not only complement community studies but also provide important information about the population where the assessment of psychopathic traits informs institutional placement, management, and release. Second, no studies have examined the stability of adolescent psychopathic using the PCL:YV. This is important given its wide use, concerns about developmental appropriateness as a downward extension of the adult PCL–R (Johnstone & Cooke, 2004), and whether the assessment protocol is sensitive to detecting changes in symptom level and severity. More important, the issue of measurement unreliability (i.e., measure and rater unreliability) has not often been addressed, as this can attenuate stability estimates.

Finally, no studies have examined developmental differences in stability across adolescence. Adolescence spans a relatively large time frame and it is likely that differences exist in early and late adolescence. In fact, there is evidence of differences with respect to a number of cognitive, emotional, and psychosocial variables, suggesting that it is important to differentiate between early, middle, and late adolescence (Steinberg & Cauffman, 1996). To address these issues and limitations of previous studies, the current study examined the 6-month stability of psychopathic traits in a sample of male adolescent offenders who were assessed with the PCL:YV, augmented to increase its sensitivity to change, and the self-report Antisocial Process Screening Device (APSD). The inclusion of both clinical ratings and self-report allows for an examination of measurement issues, given much discussion in the field about how to accurately and efficiently assess psychopathic traits. In addition, we examined stability in younger and older adolescents to investigate developmental differences in the stability of psychopathic traits.

METHOD

Participants

Participants were 83 male adolescent offenders incarcerated in minimum (31%) or maximum (69%) security youth custody centers in western Canada. Ethics approval was obtained from the Simon Fraser University Ethics Review Board prior to the start of the study. Informed consent was obtained from both participants and their parents or legal guardians. Participants were offered snacks and monetary compensation ($10) for their participation.

Participants were assessed on two occasions: once shortly after their initial recruitment (Time 1) and again after a minimum 6-month follow-up period (Time 2). Research assistants assessed 112 youth at Time 1 and 83 youth at Time 2 (i.e., an attrition rate of 26%). The most common reasons for attrition were inability to contact or locate (15 youth), geographical distance (9 youth), or refusal to participate in the follow-up (5 youth). Participants who did not complete follow-up assessments differed significantly from those who did complete assessments with respect to age and PCL:YV total scores. Those who were not assessed at follow-up were older ($M = 16.79$, $SD = 1.24$) compared to those who were assessed at follow-up ($M = 16.11$, $SD = 1.44$), $t(110) = 2.28$, $p = .02$, $d = .49$. Those who did not complete the follow-up had higher PCL:YV total scores ($M = 29.17$, $SD = 5.14$) compared to those who completed the follow-up ($M = 25.81$, $SD = 6.62$), $t(108) = 2.47$, $p = .02$, $d = .54$.

The final sample of 83 participants ranged in age from 13 to 20 years at Time 1, with a mean age of 16.11 ($SD = 1.44$). The majority were Caucasian (47%)

or Aboriginal (42%), with the remainder being of other (11%) ethnicity. Participants were incarcerated for largely violent (48%) or property (29%) offenses. The follow-up between Time 1 and Time 2 ranged from 6 to 18 months, and averaged 7.87 months ($SD = 2.47$).[1] At Time 2, participants were incarcerated in youth custody centers (66%), on youth supervision orders (31%), or incarcerated in adult institutions (2%). They ranged in age from 14 to 20 years, with a mean age of 16.73 ($SD = 1.48$). Participants were charged or convicted of largely violent (49%) or property (21%) offenses.

Although participants at Time 1 were incarcerated, whereas participants at Time 2 were either incarcerated or in the community, there was no significant difference in terms of how much time adolescents had recently spent in custody (i.e., within the past 6 months), $t(82) = 0.59$, $p = .56$. The mean number of months participants were incarcerated at Time 1 was 3.29 ($SD = 1.69$) whereas the mean number of months incarcerated at Time 2 was 3.13 ($SD = 2.18$).

Measures and Procedure

PCL:YV–Lifetime ratings.   We made lifetime ratings of psychopathic traits at Time 1 using standard administration of the PCL:YV (Forth et al., 2003). The PCL:YV is a 20-item clinical rating scale that assesses psychopathic traits in adolescents between the ages of 12 and 18. Items are rated on a 3-point scale, ranging 0 (*item does not apply*), 1 (*item applies in some respects*), and 2 (*item definitely applies*), and summed to yield a total score that can range from 0 to 40. Total scores represent the extent to which an adolescent matches the prototypical psychopath. When items are omitted due to insufficient information, scores are prorated. Numerous studies suggest the PCL:YV is a reliable and valid assessment instrument. Reviews report excellent internal consistency, ranging from .85 to .98, and excellent interrater reliability of total scores, ranging from .81 to .98 (Book, Clark, Forth, & Hare, 2006; Forth, 2005).

The PCL:YV items can also be summed to yield factor scores. The PCL:YV manual reports support for both the three- and four-factor hierarchical models. The three-factor model was developed by Cooke and Michie (2001), based primarily on analysis of the PCL–R in adults: Arrogant and Deceitful Interpersonal Style (Factor 1), Deficient Affective Experience (Factor 2), and Impulsive and Irresponsible Behavioral Style (Factor 3). The four-factor model encompasses the three

factors of Cooke and Michie (2001) but adds a fourth factor that represents the criminal behavioral features of psychopathy (Factor 4). We investigated the three-factor model to examine the core personality features of psychopathy.

In this study, Lifetime PCL:YV ratings were completed by one of three trained raters on the basis of an interview and a review of file information. PCL:YV Total scores ranged from 9 to 39, with a mean of 25.81 ($SD = 6.62$). The internal consistency (Cronbach's alpha) of Total scores was .87 and the mean interitem correlation (MIC) was .26. The mean Factor 1, Factor 2, and Factor 3 scores were 4.12 ($SD = 1.87$), 5.15 ($SD = 1.68$), and 7.20 ($SD = 1.89$), respectively. The internal consistencies of the Factor 1, Factor 2, and Factor 3 scores were .72, .66, and .75, respectively. The MICs of the Factor 1, Factor 2, and Factor 3 scores were .39, .34, and .38, respectively. Interrater reliability (ICC_1), calculated on a subset of 24 participants (29%), was excellent: .95 for Total scores, .77 for Factor 1, .77 for Factor 2, and .75 for Factor 3.

PCL:YV–6-month ratings.   We revised the PCL:YV protocol at Time 1 and Time 2 to rate psychopathic traits during the previous 6 months. This involved augmenting the standard interview protocol at Time 1 to ask youth additional questions focusing on their functioning in the past 6 months. The 6-month PCL:YV ratings at Time 1 (original) were made concurrently with the Lifetime PCL:YV ratings. The 6-month PCL:YV ratings at Time 2 were made by a rater different from the rater who completed the Time 1 assessment. Thus, 6-month PCL:YV ratings at Time 2 were blind to the Lifetime and 6-month PCL:YV ratings at Time 1. The 6-month PCL:YV ratings at Time 2 were based on an interview and review of file information. At Time 2, the interview focused specifically on the youth's functioning in the past 6 months (i.e., a standard lifetime rating was not conducted). Files reviewed at Time 2 provided comprehensive information on participants' functioning across multiple contexts (i.e., family, school, and community).

Given that the original PCL:YV 6-month ratings at Time 1 were made concurrently with the Lifetime PCL:YV ratings, we obtained an independent set of 6-month ratings based on audiotapes of the interviews and a review of file information from a second set of different raters (independent). In other words, this set of raters differed from the set of raters who completed the concurrent PCL:YV Lifetime and 6-month ratings at Time 1. This independent set of raters also completed a 6-month PCL:YV rating at Time 2. In other words, we obtained two PCL:YV 6-month ratings at both Time 1 and Time 2. One of two raters focused on participants' functioning within the 6-month time frame (i.e., did

---

[1]There were no significant differences between adolescents who had shorter (i.e., fewer than 10 months) and longer (i.e., 10 months or more) follow-up lengths in terms of mean age or PCL:YV scores at Time 1 and Time 2.

TABLE 1
Mean Time 1 and Time 2 6-Month PCL:YV Ratings by Original and
Independent Raters

PCL:YV	Original Raters			Independent Raters		
	M (SD)	α	MIC	M (SD)	α	MIC
Time 1						
Total	20.18 (7.50)	.89	.29	21.67 (8.03)	.87	.25
Factor 1	3.28 (2.11)	.76	.45	3.67 (2.46)	.79	.48
Factor 2	4.28 (1.80)	.59	.26	4.90 (2.00)	.58	.25
Factor 3	5.89 (2.30)	.76	.40	5.98 (2.47)	.67	.29
Time 2						
Total	17.28 (7.26)	.88	.27	18.73 (8.47)	.88	.27
Factor 1	3.25 (1.74)	.60	.28	3.44 (2.09)	.60	.27
Factor 2	3.65 (1.71)	.65	.31	4.39 (2.33)	.70	.37
Factor 3	5.17 (2.43)	.80	.44	5.35 (2.67)	.72	.35

Note. PCL:YV = Psychopathy Checklist: Youth Version (Forth et al., 2003); MIC = mean interitem correlation.

not conduct a Lifetime assessment) and reviewed the appropriate edited file information (i.e., file information pertaining to lifetime functioning was removed). Thus, these independent ratings were blind to the original Lifetime, Time 1, and Time 2 PCL:YV ratings.

The mean 6-month PCL:YV Total and Factor scores of the first (original) and second (independent) set of raters at Time 1 and Time 2 are presented in Table 1. As indicated in the table, the internal consistencies were generally within the acceptable range of .70 or higher (Nunnally & Bernstein, 1994). Although the MICs of the total scores were in the acceptable range of .15 to .50 for broad higher order constructs, several of the MICs of the factor scores were below the acceptable range of .40 to .50 for narrower constructs (Clark & Watson, 1995). Interrater reliability (ICC$_2$) of the 6-month ratings between the original and independent raters at Time 1 was excellent: .92 for Total scores, .89 for Factor 1, .79 for Factor 2, and .89 for Factor 3. Interrater reliability of the 6-month ratings at Time 2 was also excellent: .95 for Total scores, .92 for Factor 1, .78 for Factor 2, and .88 for Factor 3. Given that there was excellent interrater reliability between the original and independent ratings, we averaged across the 6-month PCL:YV ratings (see Table 2 for descriptive information).

To evaluate the reliability and validity of the 6-month PCL:YV ratings, we calculated the correlations between the original Lifetime and independent 6-month PCL:YV ratings at Time 1. The correlations for the Total, Factor 1, Factor 2, and Factor 3 scores were .71, .72, .58, and .58, respectively. All correlations indicated a strong association and were significant at the $p < .01$ level. Second, we examined associations between the original Lifetime and independent 6-month PCL:YV ratings at Time 1 with self-reported aggression, and official and self-reported criminal history. Self-reported

aggression was assessed using the Form-Function Aggression Measure (Little, Jones, Henrich, & Hawley, 2003), designed to assess the forms (overt, relational) and functions (instrumental, reactive) of aggression. The measure demonstrates construct validity with respect to frustration tolerance and hostility. In this study, we examined the instrumental and reactive subscales of aggression. Official previous violent, nonviolent, and total charges and convictions were obtained from a computer database. Criminal history was calculated as a proportion to account for differences in the opportunity to commit offenses. The correlations between Lifetime PCL:YV Total scores, and instrumental and reactive aggression were .34 and .24, respectively. For violent, nonviolent, and total offenses, the associations were .35, .39, and .56, respectively. The correlations between the 6-month PCL:YV Total scores, and instrumental and reactive aggression were .44 and .35, respectively. For violent, nonviolent, and total offenses, the associations were .22, .38, and .42, respectively. All correlations indicated a moderate to strong association and were significant at the $p < .05$ level. Of importance, there were no statistically significant differences in the magnitude of associations between the Lifetime and 6-month ratings. In general, these findings indicated that the augmented PCL:YV protocol yielded highly reliable 6-month ratings of psychopathic traits, the original 6-month ratings at Time 1 were not substantially influenced by knowledge of Lifetime ratings, and there was preliminary evidence for the validity of the augmented protocol with respect to aggression and criminal history.

APSD. The self-report APSD (Caputo, Frick, & Brodsky, 1999) is a 20-item measure modeled after the PCL–R to screen for psychopathic traits in adolescents.

TABLE 2
PCL:YV and APSD Descriptives at Time 1 and Time 2

PCL:YV	Time 1			Time 2			
	M (SD)	α	MIC	M (SD)	α	MIC	d
Total	20.90 (7.51)	.90	.31	18.01 (7.66)	.91	.33	.38
Factor 1	3.48 (2.18)	.84	.57	3.31 (1.84)	.69	.37	
Factor 2	4.59 (1.73)	.63	.29	3.99 (1.86)	.75	.42	.34
Factor 3	5.91 (2.28)	.77	.41	5.27 (2.28)	.81	.46	.28
APSD							
Total	16.74 (5.49)	.75	.15	15.06 (5.72)	.81	.17	.30
Narcissism	4.73 (2.76)	.75	.30	4.23 (2.53)	.71	.26	
Callous-Unemotional	2.79 (1.31)	.22	.06	2.80 (1.51)	.46	.17	
Impulsivity	5.34 (1.91)	.58	.22	4.79 (1.95)	.61	.25	.28

Note. PCL:YV = Psychopathy Checklist: Youth Version (Forth et al., 2003); APSD = self-report Antisocial Process Screening Device (Caputo et al., 1999); MIC = mean interitem correlation; d = Cohen's d effect size for the effect of time.

Respondents indicate the degree to which they believe the statements generally reflect their personality using a 3-point scale, ranging 0 (*not at all true*), 1 (*sometimes true*), and 2 (*definitely true*). When participants omit items, scores are prorated. Items can be summed to yield total scores that range from 0 to 40 and three factor scores (Vitacco, Rogers, & Neumann, 2003): Narcissism (Factor 1), Callous-Unemotional (Factor 2), and Impulsivity (Factor 3). As recommended by Poythress et al. (2006), we omitted Items 19 and 20 when calculating scores on the Callous-Unemotional factor. The self-report APSD total and factor scores demonstrate construct validity as evidenced by associations with violence, recidivism, and program noncompliance (Falkenbach, Poythress, & Heide, 2003; Spain, Douglas, Poythress, & Epstein, 2004). Participants completed the self-report APSD at the end of their interview at Time 1 and Time 2. Time 1 APSD Total scores ranged from 3 to 28, with a mean of 16.74 ($SD = 5.49$). Time 2 APSD Total scores ranged from 1 to 26, with a mean of 15.06 ($SD = 5.72$). The means, internal consistencies, and MICs of the Total and Factor scores at Time 1 and Time 2 are reported in Table 2.

Analyses: Generalizability Theory

When examining the stability of a construct, it is important to account for the reliability of the measure as this may attenuate the magnitude of stability coefficients. Stability estimates indexed with Pearson and ICCs include error due to measurement unreliability and therefore provide conservative estimates. A more accurate assessment of stability may be obtained with generalizability theory (G theory), an extension of classical test theory that recognizes multiple sources of measurement variance that typically are subsumed under a single error term (Marcoulides, 2000; Shavelson & Webb, 1991). The variance attributable to different sources of variability or error are provided and estimated simultaneously in a single analysis. Any potential source of error is referred to as a facet, and the levels within each facet are referred to as conditions. For example, facets for the PCL:YV include items, raters, and time (i.e., Time 1 and Time 2), whereas facets for the APSD include items and time. Both measures are composed of 20 items; therefore, the item facet has 20 conditions.

In G theory analyses, reliability or stability is indexed by a G coefficient (Marcoulides, 2000), which represents the reliability of scores across items, raters, or assessments, and can range from 0 to 1. As a G coefficient is a form of intraclass correlation, similar guidelines may assist in interpreting G coefficients: .75 and above as high, .60 to .74 as moderate, and .59 and below as low (Cicchetti & Sparrow, 1981). The following sources of variability are noted: variation in the degree to which the items assess the construct (Item), variation in the degree to which scores remain consistent between the two assessments when averaged across participants and items (Time), variation in the rank ordering of participants across the items (Participant × Item interaction), variation in the rank ordering of participants between the two assessments (Participant × Time interaction), and variation in the degree to which the items assess the construct between the two assessments (Item × Time interaction). G theory analyses were conducted with a program developed by Mushquash and O'Connor (2006).

RESULTS

Temporal Stability of 6-month PCL:YV Ratings

Stability of total and factor scores. We conducted a series of G theory analyses to examine the stability of the 6-month PCL:YV Total and Factor scores across Time 1 and Time 2. The facets in this design included items, raters, and time. However, because raters were nested within time (i.e., all raters did not assess all participants at both assessments), the variance attributable to raters could not be estimated separately. Given excellent interrater reliability at Time 1 ($ICC_2 = .79$–.92) and Time 2 ($ICC_2 = .78$–.95) between the original and independent assessments, ratings were averaged and analyses conducted to examine stability. Therefore, the final design included items and time as facets.

For the 6-month Total and Factor ratings, there was little variance attributable to Item (range = .02–.07), Time (range = .00–.01), the Participant × Item interaction (range = .01–.06), the Participant × Time interaction (range = .04–.07), and the Item × Time interaction (range = .00–.01). In other words, there was little variation in the extent to which the items assessed the construct and in the rank ordering of participants on the items across time. Furthermore, scores were consistent between Time 1 and Time 2, there was little variability in the rank ordering of scores across time, and the items assessed psychopathic traits similarly at both assessments. The G coefficients (with 95% confidence intervals) for the Total, Factor 1, Factor 2, and Factor 3 scores were .75 (.71–.79), .72 (.67–.77), .49 (.43–.55), and .70 (.65–.75), respectively.[2] These results indicated high 6-month

---

[2]Analyses were rerun omitting 5 participants who were older than age 18 and those who had a longer (i.e., 10 months or more) follow-up length. There were no substantial differences in G coefficients. Analyses were also run separately for those with recent incarceration lengths of 3 months or fewer and those longer than 3 months. There were no substantial differences in G coefficients, with the exception of Factor 3, which demonstrated greater stability in adolescents who were recently incarcerated for more than 3 months.

stability of PCL:YV Total scores, when averaged across participants and items; they also indicated that unreliability because of raters and items were not substantial sources of error. In contrast, the interpersonal and behavioral factors evidenced moderate stability whereas the affective traits evidenced low stability.

Magnitude and direction of change. A repeated measures multivariate analysis of variance was conducted to examine mean-level changes in scores (see Table 2). There were statistically significant decreases in 6-month PCL:YV total, affective, and behavioral scores, although the effect sizes indicated the differences were not substantial. To provide a more detailed context of change, we examined the proportion of adolescents whose 6-month PCL:YV ratings increased, decreased, and remained the same. Using the standard error of measurement (SEM) of Lifetime PCL:YV Total scores as a guideline (Forth et al., 2003), scores within one SEM across Time 1 and Time 2 (i.e., total scores within ±3) were defined as "no change." Using this criterion, approximately one third of the sample exhibited no change in scores (34%). There were relatively few participants whose scores increased substantially; 6% of the sample had total scores that increased by two SEMs and 10% that increased by three or more SEMs. In contrast, scores decreased for a substantial proportion of participants; 26% showed a decrease in total scores by two SEMs and 23% demonstrated a decrease of three or more SEMs.

Temporal Stability of APSD Ratings

Stability of total and factor scores. A series of G theory analyses was conducted to examine the stability of the self-report APSD Total and Factor scores. The facets in this design included items and time. For both the Total and Factor scores, there was little variance attributable to Item (range = .04–.14), Time (range = .00–.01), the Participant × Item interaction (range = .06–.10), and the Participant × Time interaction (range = .00–.04), and no variance attributable to the Item × Time interaction. In other words, there was little variation in the extent to which the items assessed the construct, and in the rank ordering of participants on the items across time. Furthermore, scores were similar across Time 1 and Time 2, there was little variability in the rank ordering of scores across time, and the items assessed psychopathic traits similarly at both assessments. The G coefficients (with 95% confidence intervals) for the Total, Narcissism, Callous-Unemotional, and Impulsivity scores were .72 (.69–.75), .60 (.55–.65), .48 (.43–.53), and .62 (.57–.67), respectively. The results

suggested moderate 6-month stability of self-report APSD Total scores, when averaged across participants and items, and that rater and item unreliability were not substantial sources of error. Similarly, the narcissistic and impulsive traits evidenced moderate stability whereas the callous-unemotional traits demonstrated low stability.

Magnitude and direction of change. A repeated measures multivariate analysis of variance indicated there was a statistically significant decrease in APSD Total and Impulsivity scores over 6 months, although the effect sizes suggested these differences were not substantial (see Table 2). Similar to the analyses for the PCL:YV, the proportion of adolescents whose scores increased, decreased, and remained the same was calculated. Using the SEM of APSD Total scores as a guideline (i.e., total scores within ±3; Frick & Hare, 2001), 54% of the sample exhibited no change in scores across Time 1 and Time 2. There were relatively few participants whose scores increased substantially; 10% of the sample had total scores that increased by two SEMs and 4% that increased by three or more SEMs. In contrast, 13% showed a decrease in total scores by two SEMs and 20% demonstrated a decrease of three or more SEMs.

Comparing Temporal Stability of Ratings at Different Developmental Stages

We took a median split of the sample to evaluate the temporal stability of psychopathic traits across developmental stages. This yielded 34 younger adolescent offenders (aged 14–16) and 49 older adolescent offenders (aged 17–20). Of importance, there were no significant differences between younger and older adolescents with respect to the mean number of previous violent, nonviolent, and supervision offenses, PCL:YV and APSD Total and Factor scores at baseline, follow-up length, and recent length of incarceration at follow-up. The G coefficients (with 95% confidence intervals) for the 6-month PCL:YV Total, Factor 1, Factor 2, and Factor 3 scores were .57 (.51–.63), .63 (.54–.72), .33 (.25–.41), and .53 (.48–.58), respectively, in younger adolescents, and .79 (.74–.84), .77 (.71–.83), .60 (.53–.67), and .75 (.68–.82), respectively, in older adolescents. These results indicated low to moderate stability of psychopathic traits in younger adolescents, whereas there was moderate to high stability in older adolescents. The G coefficients (with 95% confidence intervals) for the APSD Total, Narcissism, Callous-Unemotional, and Impulsivity scores were .67 (.62–.72), .60 (.54–.66), .45 (.38–.52), and .65 (.58–.72), respectively, in younger adolescents, and .74 (.70–.78), .59 (.52–.66), .50 (.43–.57),

and .58 (.51–.65), respectively, in older adolescents. These results indicated low to moderate stability in both younger and older adolescents. In general, there was greater stability of psychopathic traits in older adolescents when assessed with the PCL:YV. In contrast, stability was comparable between younger and older adolescents using the APSD, with the exception of impulsivity, which demonstrated greater stability in younger adolescents relative to older adolescents.

DISCUSSION

Although researchers have started to investigate the stability of child and adolescent psychopathic traits (e.g., Frick et al., 2003; Loney et al., 2007; Lynam et al., 2007), the issue has not been thoroughly examined in clinical and forensic populations, which may limit our understanding of stability in adolescence. More important, there are concerns that the PCL:YV is not sensitive to detecting changes in psychopathic traits across time. We augmented the PCL:YV protocol to increase its sensitivity to detecting change and examined the 6-month stability of psychopathic traits in a sample of adolescent offenders using two different methodological approaches (clinical ratings vs. self-report).

Stability of Psychopathic Traits

The current study is the first to address concerns about the ability of the PCL:YV to detect changes in adolescent psychopathic traits across time. Our attempt to augment the protocol appears promising. More specifically, there was good to excellent interrater reliability and internal consistency, and strong associations between the original and augmented protocols. Furthermore, associations with aggression and criminality provided preliminary evidence for the validity of the protocol. However, the findings and interpretations should be viewed with caution given methodological limitations. Psychopathic traits were assessed within a specific time frame, which may not be the optimal method of assessment. As such, our findings likely provide a conservative estimate of the short-term stability of psychopathic traits in adolescent offenders.

With this caveat in mind, this unique approach for assessing adolescent psychopathic traits contributes to the growing literature on the stability of child and adolescent psychopathic traits in several ways. First, we found moderate to high stability of psychopathic traits, as indexed by PCL:YV and APSD total scores, which is consistent with previous studies (e.g., Loney et al., 2007; Lynam et al., 2007) and research on general personality traits (e.g., Roberts & Delvecchio, 2000). At the factor level, the interpersonal, affective, and

behavioral traits demonstrated low to moderate stability. Of interest, the APSD stability estimates were lower than those reported by Frick et al. (2003), which may reflect sample (i.e., normative children vs. adolescent offenders) or methodological (i.e., parent- vs. self-report) differences.

Second, changes in the magnitude and direction of PCL:YV and APSD scores across time supported the findings of moderate to high stability. There were significant mean-level changes in PCL:YV total, affective, and behavioral traits, and in APSD total and impulsivity traits, although the differences were not substantial. Consistent with Frick et al. (2003), decreases were more likely than increases in scores, with approximately one fifth to one fourth of participants' scores demonstrating substantial decreases across assessments. This is particularly interesting given that we employed a different method for examining the nature of stability. However, the decreases observed in PCL:YV scores may reflect limitations of assessing psychopathic traits within a specific time frame.

Third, findings were consistent between the PCL:YV and APSD with respect to the relative stability of the different symptom clusters. Both the PCL:YV and APSD demonstrated greater stability of the interpersonal and behavioral traits relative to the affective traits. This may be because interpersonal and behavioral traits are more easily observed through concrete indicators. However, the lower relative stability of the affective traits should be viewed with caution as this may reflect measurement issues. The interrater reliability and internal consistencies were less than optimal, suggesting that it may be difficult to assess affective deficits, such as the lack of emotional attachment or remorse, over brief periods.

Finally, there were developmental differences in stability whereby psychopathic traits evidenced greater stability in older adolescents (i.e., 17–20 years) than younger adolescents (i.e., 14–16 years). The finding is further supported in that there was no overlap in the range of stability coefficients, although this held true only for the PCL:YV. However, these findings should be viewed with caution given the small sample size and short time frame for follow-up. Although these findings are consistent with evidence of differences across early, middle, and late adolescence with respect to a number of psychosocial variables (e.g., Steinberg & Cauffman, 1996), it may also suggest that the PCL:YV is less reliable and valid for assessing psychopathic traits in younger adolescents.

Of interest, our results are inconsistent with previous studies that have found moderate to high stability of interpersonal callousness across adolescence (e.g., Obradovic et al., 2007; Pardini & Loeber, 2008). One reason may be because of methodology, in that previous

studies examined parent and teaching ratings whereas the current study examined clinical ratings and self-report. However, Pardini and Loeber found significant variability in individual trajectories across adolescence. Taken together, these findings suggest some potential developmental differences and highlight the need to examine factors that may account for developmental differences in stability.

Comparison of the PCL:YV and the APSD

Similar stability estimates were found between the PCL:YV and self-report APSD total scores. However, at the symptom cluster level, the interpersonal, affective, and behavioral traits evidenced greater stability when assessed by the PCL:YV relative to the APSD. Of interest, despite following the recommendations of Poythress et al. (2006) to improve the internal consistency of the APSD Affective subscale, the internal consistency in this sample remained low. As such, it may be difficult to assess callous-unemotional traits through self-report in adolescent offenders. In general, differences between the PCL:YV and APSD may be due to methodology (i.e., clinical judgment vs. self-report) in that the APSD may not adequately capture psychopathic traits in a manner analogous to the PCL:YV (Lee, Vincent, Hart, & Corrado, 2003; Murrie & Cornell, 2002). To illustrate, there are differences with respect to the items that load onto the specific factors. Although APSD factors are intended to correspond with the PCL:YV factors, not all parallel items of the APSD load onto the same PCL:YV factors. For example, the PCL:YV item Shallow Affect loads onto the affective factor whereas the parallel APSD item ("My emotions are shallow and fake") loads onto the interpersonal factor. In addition, the APSD attempts to capture analogous PCL:YV items with a simple, one-sentence statement that may not encompass the complexity of the PCL:YV descriptions.

Limitations

The following limitations should be noted. First, despite promising evidence for the augmented protocol in detecting changes in psychopathic traits, questions remain about reliability and validity. We attempted to address the reliability of assessments by obtaining independent ratings. However, an important question is whether assessments of psychopathic traits completed within a specific time frame demonstrate associations with other constructs (e.g., recidivism, psychopathology) in a similar manner to ratings conducted using the standard protocol. Stronger evidence for the validity of the protocol will assist in resolving the debate as to whether the PCL-R and PCL:YV protocols are sensitive to detecting change. Future studies should examine whether other factors influence the reliability and validity of ratings, such as the amount of time detained during follow-up and varying the time frame. These results may then provide important information about how best to detect changes in psychopathic traits.

Second, our findings address stability in incarcerated male adolescent offenders. Although there is no theoretical reason to believe that stability should be moderated by gender, at the very least, this is an important question for generalizability. Furthermore, adolescents likely participated in various interventions during incarceration and supervision, although there was no reason to believe there were any systematic group differences in the types of interventions administered. Future studies should examine whether interventions influence the manifestation and stability of psychopathic traits. Finally, there was selective attrition of older adolescents and adolescents with higher PCL:YV scores, although examination of the effect sizes and absolute magnitude of the differences suggests these were not substantial. Despite this, the inclusion of these youth may have resulted in higher stability estimates on the assumption that psychopathic traits are more crystallized in older adolescents, and therefore more likely to exhibit stability.

Implications for Research, Policy, and Practice

Our findings also have important clinical and policy implications for understanding the construct of psychopathy in adolescents. Although adolescent psychopathic traits are associated with less responsivity to treatment (e.g., O'Neill et al., 2003), recent evidence indicates that intensive treatment with psychopathic juvenile offenders can lead to reductions in violent recidivism (e.g., Caldwell et al., 2006). Similarly, our findings of moderate to high stability suggest that interventions need to be intense if the goal is to reduce adolescent psychopathic traits and prevent the development of the disorder. More specifically, interventions should be administered in early adolescence when the traits appear more malleable.

Second, our findings of low to moderate stability of the interpersonal, affective, and behavioral symptoms highlight the need to further our understanding of how these dimensions contribute to the disorder. In other words, more refined analyses examining their construct and predictive validity, and whether different symptom interactions are associated with differential outcomes. Third, multiple assessments across adolescence will provide a better understanding of the progression and stability of psychopathic traits during adolescence. It is important to note that stability coefficients do not specify what variables or processes influence (in)stability. For example, Frick and Dantagnan (2005) found that

more life stressors were associated with greater stability of conduct problems in children with conduct problems and callous-unemotional traits, whereas greater association with delinquent peers was associated with less stability. In contrast, Barry et al. (2008) found that higher social competence was associated with decreases in narcissism.

Finally, evidence of potential developmental differences in stability underscore the importance of early identification of emerging psychopathic traits to prevent the negative outcomes associated with the disorder. Although the total scores demonstrated moderate to high stability, we currently do not have a comprehensive understanding of the developmental progression of psychopathic traits, what factors or mechanisms are involved in the manifestation of traits, or appropriate guidelines for the purposes of diagnosis. Therefore, it would be professionally and ethically inappropriate to use specific scores to categorize children and adolescents for clinical and legal decisions (e.g., program suitability, transfer to adult court). In sum, accurately assessing child and adolescent psychopathic traits will allow us to better track the causes, correlates, and developmental progression of the disorder, and identify theoretically meaningful moderators and developmental periods associated with (in)stability.

REFERENCES

Alterman, A. I., Cacciola, J. S., & Rutherford, M. J. (1993). Reliability of the revised Psychopathy Checklist in substance abuse patients. *Psychological Assessment*, *5*, 442–448.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Barry, T. D., Barry, C. T., Deming, A. M., & Lochman, J. E. (2008). Stability of psychopathic characteristics in childhood: The influence of social relationships. *Criminal Justice and Behavior*, *35*, 244–262.

Bazana, P. G., & Stelmack, R. M. (2004). Stability of personality across the life span: A meta-analysis. In R. M. Stelmack (Ed.), *On the psychobiology of personality: Essays in honor of Marvin Zuckerman* (pp. 113–144). New York: Elsevier Science.

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

Book, A. S., Clark, H. J., Forth, A. E., & Hare, R. D. (2006). The Psychopathy Checklist-Revised and the Psychopathy Checklist: Youth Version. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 147–179). Mahwah, NJ: Erlbaum.

Burke, J. D., Loeber, R., & Lahey, B. B. (2007). Adolescent conduct disorder and interpersonal callousness as predictors of psychopathy in young adults. *Journal of Clinical Child ad Adolescent Psychology*, *36*, 334–346.

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

Caputo, A. A., Frick, P. J., & Brodsky, S. L. (1999). Family violence and juvenile sex offending: The potential mediating role of psychopathic traits and negative attitudes toward women. *Criminal Justice and Behavior*, *26*, 338–356.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behaviour. *American Journal of Mental Deficiency*, *86*, 127–137.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309–319.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, *13*, 171–188.

Douglas, K. S., Vincent, G. M., & Edens, J. F. (2006). Risk for criminal recidivism: The role of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 533–554). New York: Guilford.

Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth psychopathy and criminal recidivism: A meta-analysis of the Psychopathy Checklist measures. *Law and Human Behavior*, *31*, 53–75.

Edens, J. F., Skeem, J. L., Cruise, K. R., & Cauffman, E. (2001). Assessment of "juvenile psychopathy" and its association with violence: A critical review. *Behavioral Sciences and the Law*, *19*, 53–80.

Falkenbach, D. M., Poythress, N. G., & Heide, K. M. (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.

Forth, A. E. (2005). Hare Psychopathy Checklist: Youth Version. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 324–338). New York: Guilford.

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

Frick, P. J. (2002). Juvenile psychopathy from a developmental perspective: Implications for construct development and use in forensic assessments. *Law and Human Behavior*, *26*, 247–253.

Frick, P. J., & Dantagnan, A. L. (2005). Predicting the stability of conduct problems in children with and without callous–unemotional traits. *Journal of Child and Family Studies*, *14*, 469–485.

Frick, P. J., & Hare, R. D. (2001). *Antisocial Process Screening Device (APSD): Technical manual*. Toronto, Canada: Multi-Health Systems.

Frick, P. J., Kimonis, E. R., Dandreaux, D. M., & Farell, J. M. (2003). The 4 year stability of psychopathic traits in non-referred youth. *Behavioral Sciences and the Law*, *21*, 713–736.

Gretton, H. M., Hare, R. D., & Catchpole, R. E. H. (2004). Psychopathy and offending from adolescence to adulthood: A 10-year follow-up. *Journal of Consulting and Clinical Psychology*, *72*, 636–645.

Hare, R. D. (1991). *Manual for the hare psychopathy checklist–revised*. Toronto, Canada: Multi-Health Systems.

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

Harris, G. T., & Rice, M. E. (2006). Treatment of psychopathy: A review of empirical findings. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 555–572). New York: Guilford.

Hart, S. D., Watt, K. A., & Vincent, G. M. (2002). Commentary on Seagrave and Grisso: Impressions of the state of the art. *Law and Human Behavior*, *26*, 241–245.

Hawes, D. J., & Dadds, M. R. (2007). Stability and malleability of callous–unemotional traits during treatment for childhood conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *36*, 347–355.

Johnson, J. G., Cohen, P., Kasen, S., Skodol, A. E., Hamagami, F., & Brook, J. S. (2000). Age-related change in personality disorder trait

levels between early adolescence and adulthood: A community-based longitudinal investigation. *Acta Psychiatrica Scandinavica*, *102*, 265–275.

Johnstone, L., & Cooke, D. J. (2004). Psychopathic-like traits in childhood: Conceptual and measurement concerns. *Behavioral Sciences and the Law*, *22*, 103–125.

Lee, Z., Vincent, G. M., Hart, S. D., & Corrado, R. R. (2003). The validity of the Antisocial Process Screening Device as a self-report measure of psychopathy in adolescent offenders. *Behavioral Sciences and the Law*, *21*, 771–786.

Little, T. D., Jones, S. M., Henrich, C. C., & Hawley, P. H. (2003). Disentangling the "whys" from the "whats" of aggressive behaviour. *International Journal of Behavioral Development*, *27*, 122–133.

Loney, B. R., Taylor, J., Butler, M. A., & Iacono, W. G. (2007). Adolescent psychopathy features: 6-year temporal stability and the prediction of externalizing symptoms during the transition to adulthood. *Aggressive Behavior*, *33*, 242–252.

Lynam, D. R. (2002). Fledgling psychopathy: A view from personality theory. *Law and Human Behavior*, *26*, 255–259.

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

Marcoulides, G. A. (2000). Generalizability theory. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 527–551). San Diego, CA: Academic Press.

Mattanah, J. J. F., Becker, D. F., Levy, K. N., Edell, W. S., & McGlashan, T. H. (1995). Diagnostic stability in adolescents followed up 2 years after hospitalization. *American Journal of Psychiatry*, *152*, 889–894.

Murrie, D. C., & Cornell, D. G. (2002). Psychopathy screening of incarcerated juveniles: A comparison of measures. *Psychological Assessment*, *14*, 390–396.

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

Mushquash, C., & O'Connor, B. (2006). SPSS and SAS programs for generalizability theory analyses. *Behavior Research Methods*, *38*, 542–547.

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Obradovic, J., Pardini, D. A., Long, J. D., & Loeber, R. (2007). Measuring interpersonal callousness in boys from childhood to adolescence: An examination of longitudinal invariance and temporal stability. *Journal of Clinical Child and Adolescent Psychology*, *36*, 276–292.

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.

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

Pardini, D. A., & Loeber, R. (2008). Interpersonal callousness trajectories across adolescence: Early social influences and adult outcomes. *Criminal Justice and Behavior*, *35*, 173–196.

Poythress, N. G., Douglas, K. S., Falkenbach, D., Cruise, K., Lee, Z., Murrie, D. C., et al. (2006). Internal consistency reliability of the self-report Antisocial Process Screening Device. *Assessment*, *13*, 107–113.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*, 3–25.

Rutherford, M., Cacciola, J. S., Alterman, A. I., McKay, J. R., & Cook, T. G. (1999). The 2-year test–retest reliability of the Psychopathy Checklist–Revised in methadone patients. *Assessment*, *6*, 285–291.

Salekin, R. T., Neumann, C. S., Leistico, A. R., DiCicco, T. M., & Duros, R. L. (2004). Psychopathy and comorbidity in a young offender sample: Taking a closer look at psychopathy's potential importance over disruptive behavior disorders. *Journal of Abnormal Psychology*, *113*, 416–427.

Schroeder, M. L., Schroeder, K. G., & Hare, R. D. (1983). Generalizability of a checklist for assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *51*, 511–516.

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

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.

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

Steinberg, L., & Cauffman, E. (1996). Maturity of judgment in adolescence: Psychosocial factors in adolescent decision making. *Law and Human Behavior*, *20*, 249–272.

Vitacco, M. J., Rogers, R., & Neumann, C. S. (2003). The Antisocial Process Screening Device: An examination of its construct and criterion-related validity. *Assessment*, *10*, 143–150.

Westen, D., & Chang, C. (2000). Personality pathology in adolescence: A review. In A. H. Esman, L. T. Flaherty, & H. A. Horowitz (Eds.), *Adolescent psychiatry: Developmental and clinical studies, Volume 25* (pp. 61–100). Hillsdale, NJ: The Analytic Press.

World Health Organization. (2000). *International classification of diseases, tenth revision*. Geneva, Switzerland: Author.