Identifying Persistently Antisocial Offenders Using the Hare Psychopathy Checklist and DSM Antisocial Personality Disorder Criteria

Tracey A. Skilling
Queen’s University and Penetanguishene Mental Health Centre

Grant T. Harris and Marnie E. Rice
Penetanguishene Mental Health Centre

Vernon L. Quinsey
Queen’s University

Early starting, lifetime criminal persistence has been called sociopathy, antisocial personality disorder, and psychopathy. There is, however, disagreement about its core features and which measure is best for identifying such individuals. In the 1st of 2 studies of male offenders (n = 74), we found a large association between scores on the Psychopathy Checklist—Revised (PCL–R; R. D. Hare, 1991) and the Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM–IV]; American Psychiatric Association, 1994) antisocial personality disorder criteria scored as a scale. The second study (n = 684) replicated this finding and found that, as previously shown for PCL–R scores, a discrete class (or taxon) also underlies scores on items reflecting antisocial personality disorder. The high association among these sets of items and their similarity in predicting violence suggested that the same natural class underlies each. Results indicated that life-course-persistent antisociality can be assessed well by measures of psychopathy and antisocial personality disorder.

The most persistent adult male offenders (5%–6% of offender populations) are responsible for a disproportionate (about 50%) amount of crime (see Farrington, Ohlin, & Wilson, 1986, for a review; see also Hare & Jutai, 1983; Kosson, Smith, & Newman, 1990). They are also more difficult to manage in correctional and institutional settings (Forth, Hart, & Hare, 1990; Hare & McPherson, 1984; Wong, 1984). They reoffend and violate conditions of release sooner and more often, and are at higher risk to reoffend violently, than other offenders (Hare, 1981; Hare & McPherson, 1984; Harris, Rice, & Cormier, 1991; Hart, Kropp, & Hare, 1988; Rice, Harris, & Quinsey, 1990; Serin, 1991; Serin, Peters, & Barbaree, 1990; Wong, 1984). In addition, these offenders begin their criminal careers earlier and persist in them longer (e.g., Harris et al., 1991; Wolfgang, Thornberry, & Figlio, 1987).

Clearly, the identification of these highest risk offenders depends on two things: early age of onset of violent and antisocial conduct and persistence of that antisocial and criminal behavior. Moffitt (1993) proposed that, during adolescence, delinquent behavior is carried out by two distinct groups: a majority whose antisocial behavior is confined to adolescence and a smaller group who begin earlier and continue in antisocial behavior through adulthood. Variations in this taxonomy of adolescent-limited and life-course-persistent antisocial behavior have been proposed by other researchers as well (e.g., Loeber, 1982; Patterson, 1992) and many researchers have commented on the remarkable continuity of serious antisocial behavior in different samples (Loeber, 1982; Loeber & Farrington, 1998; Robins, 1978). Indeed, age of onset of behavior problems is the single best predictor of adult criminal outcomes (Farrington et al., 1990).

After age of onset and persistence of antisocial conduct, however, it remains an open and debated question as to what further characteristics, behaviors, or traits would aid in the identification of these persistently antisocial individuals. Lifelong persistent criminals have been said to have antisocial personality disorder (APD) or to be psychopathic. A requirement for a diagnosis of APD (American Psychiatric Association [APA] 1994) is evidence of conduct disorder in childhood. Similarly, the most widely used instrument for the assessment of psychopathy (the Psychopathy Checklist—Revised [PCL–R]; Hare, 1991) refers to early behavior problems and juvenile delinquency. The terms antisocial person-
ality disorder and psychopathy have often been used interchangeably to refer to this group of persistently antisocial offenders; however, researchers studying the two constructs have approached assessment differently.

Cleckley (1941, 1976) described psychopaths as superficially charming but lacking remorse, shame, and sincerity. Many scholars believe that these characteristics allow psychopaths to violate social norms, causing harm and distress to others, without regret or guilt. Expanding on Cleckley’s core characteristics, Hare developed the PCL–R to assess the extent to which someone “matches the prototypical psychopath” (Hare, 1991, p. 17). Many PCL–R items represent the central affective and interpersonal traits put forth by Cleckley, whereas other items describe an impulsive, exploitative, and antisocial lifestyle (Hare et al., 1990; Harpur, Hakstian, & Hare, 1988). Item response analyses of PCL–R scores suggested that both types of items are equally useful in measuring the extent to which someone matches the prototypical psychopath, but the interpersonal and affective items are more likely to be endorsed among offenders who are high scorers on the PCL–R (Cooke & Michie, 1997, 1999; Cooke, Michie, Hart, & Hare, 1999). The latter studies also indicated that a satisfactory fit of observed PCL–R item intercorrelations could be given by a single principal component.

It has been argued, on the other hand, that the assessment of affective or interpersonal traits used in the PCL–R requires unnecessary inference and that the phenomenon (persistent antisociality) could be more concretely and behaviorally measured (Robins, 1978). This has led to a less inferential approach to classification in the diagnosis of APD. The features described in the second edition of APA’s Diagnostic and Statistical Manual of Mental Disorders (DSM; APA, 1968) were consistent with Cleckley’s (1941, 1976) description of psychopathy but provided no explicit diagnostic criteria, describing individuals who were undersocialized, selfish, callous, lacking in guilt, irresponsible, impulsive, and failing to learn from experience. Radical changes to the criteria were introduced in the third edition (DSM–III; APA, 1980). In this edition, explicit criteria were provided in light of research on conduct disorder and these criteria emphasized readily documented rule-violating behavior in line with the recommendations of Robins (1978). The revised third edition (DSM–III–R; APA, 1987) and DSM–IV (APA, 1994) made minor changes to these behavioral criteria, and DSM–IV stated that psychopathy, sociopathy, and dissocial personality disorder were synonyms for antisocial personality. The accompanying discussion referred to other traits associated with the construct (glibness and superficial charm, as examples) but gave no guidelines on how to incorporate them in diagnosis.

Hare and others have argued that the definition of APD by DSM (DSM–III and onwards) departs markedly from the general agreement in the literature (and clinical practice) that affective characteristics are central to the construct (Harpur, Hart, & Hare, 1994; Rogers, Duncan, Lynett, & Sewell, 1994). The affective, interpersonal approach has emphasized a distinction between psychopathic personality and antisocial behavior in the absence of the constellation of affective and interpersonal features. Studies by Hare (1983, 1991, 1996) showed that dichotomous APD diagnoses (by DSM–III and DSM–III–R criteria), Cleckley’s (1941, 1976) criteria, and the PCL–R are interrelated, but that the APD criteria are overinclusive. That is, almost all participants identified as psycho-
We examined the relationship between the PCL–R and DSM APD criteria when both were treated as continuous measures and scored in a similar manner. Rutherford, Alterman, Cacciola, and McKay (1998) recently examined this question and found correlations between the two measures between .40 and .64. The present study differs from their study in important ways. Participants in the Rutherford et al. study were male and female substance abusers, rather than serious adult male offenders. In the present studies, APD criteria were scored identically to the PCL–R items (0 = absent, 1 = some indication criterion is present, or 2 = present), then summed to form a scale. It is unclear whether this scoring method or a symptom count was used by Rutherford et al.

Our next goal was to evaluate whether DSM APD criteria gave the same evidence of an underlying discrete class that we had earlier reported (using the same participants) for PCL–R items and other measures of childhood antisociality (Harris, Rice, & Quinsey, 1994). Our goal was, again without reference to cut-off scores, to ascertain the actual practical consequences of the two predominant approaches to the identification of lifetime persistent antisociality, and whether DSM APD criteria identified substantially the same discrete class of offenders that we have previously reported using PCL–R items. Finally, we examined which indicators might best detect this discrete class associated with lifetime persistent antisociality. Results have implications for both for the understanding and optimal measurement of this construct.

Study 1

As discussed earlier, the correlation between Hare’s (1991) PCL–R and the various DSM criteria for a diagnosis of APD is sizeable, but the two are not interchangeable. Many PCL–R items describe a chronically unstable and antisocial lifestyle and closely resemble the behavioral criteria of DSM–IV APD. The DSM assessment of APD is explicitly categorical. For the purposes of arriving at a diagnosis, data are dichotomized to produce present versus absent ratings, and a diagnosis of APD is rendered when a specified number of criteria are met. In order to determine, in a manner free from the statistical attenuation introduced by dichotomization at the diagnostic level, whether there are material and practical differences between the identification of persistent antisociality using the PCL–R compared to the use of APD criteria from the DSM, data gathered using both approaches need to be treated in the same way. Because both measures are composed of several items, each can be treated as a multi-item scale. If the two measures were found to measure substantially the same construct, then the empirical results obtained from both measures could be integrated, considerably simplifying the literature in this area. We hypothesized that, when the data from the two measures were treated similarly and the effects of attenuation eliminated, the two approaches would identify the same individuals. Therefore, we predicted that the agreement coefficient between the PCL–R and DSM APD criteria would be close to unity.

Method

Participants

Data were obtained from 74 male inmates from one minimum- and two medium-security Canadian federal penitentiaries. Each inmate had volunteered to participate in a previous research project. Permission was granted by the directors of the correctional service to reexamine their institutional files to code new information for the present study. Although informed consent had been given for the original study, it was (with the approval of the ethics committee) not obtained again to code the new information. Age ranged from 22 to 66 years with a mean of 37.05 years (SD = 10.70). All participants were at least 15 years old at the time of the offense that led to their inclusion in the present study and there was no recorded indication of mental illness for any participant.

Measures and Procedure

The revised 20-item version of the PCL–R (Hare, 1991) was used to assess psychopathy. Each of these assessments was based on a structured interview and an extensive review of the participant’s institutional files. Each PCL–R item was coded 0 if not present, 1 if there were some indications that the item was applicable, and 2 if it was definitely applicable. When individual items could not be coded (<5%), total scores were prorated according to the manual. These assessments were conducted by clinicians prior to and completely independently from the present study.

DSM–IV APD criteria were coded by Tracey A. Skilling retrospectively and exclusively from descriptive material contained in the participants’ psychology, case management, and sentence administration files. These institutional files included information from a variety of sources, such as information from other institutions, police reports, and psychological test reports. Most of the psychological test reports available in the files were based on clinical interviews with the participants during their prison terms. The coding of APD using DSM–IV criteria was applied to file data available at the time of data collection, and was not based on diagnoses made by consulting psychiatrists.

In addition to being blind to psychiatric opinion, PCL–R scores and institutional reports that summarized PCL–R assessments were removed from the files used for coding so that the APD assessments were performed blind to PCL–R scores. As with the PCL–R items, each APD criterion was coded 0, 1, or 2. A score of 0 was given if there was no indication that the participant had engaged in behaviors related to the criteria in question. For example, a score of 0 would be given for the conduct disorder item if there was no evidence of conduct disordered behavior in the files before Age 15. A score of 1 would be given if there were some indication that the participant had engaged in behaviors related to the specific item. For example a score of 1 would be given if one or two conduct disorder symptoms were present. And finally, a score of 2 was given if the item was clearly met (e.g., if there was evidence of three or more conduct disorder symptoms before the age of 15 years). A detailed copy of the scoring scheme is available from Tracey A. Skilling. When individual items could not be coded (<5%), total scores were prorated in keeping with the procedures used for calculating total scores on the PCL–R.

Results

Reliability

Fifteen participants were randomly chosen for an intrarater reliability check on the APD criteria, resulting in a Pearson correlation coefficient for continuous scoring of .86 and a kappa for categorical scoring of .71. No reliability data were available for the scoring of the PCL–R for the participants in this study. Measures of internal consistency (raw item alphas) as well as descriptive statistics for all scales are presented in Table 1. The alpha coefficient for the PCL–R was .92. The alpha coefficient for the APD scale was .75. These alpha coefficients are within acceptable ranges. Using the Spearman-Brown formula to predict the effect of adding APD criteria to produce a test equal in length to the PCL–R, yielded a projected alpha coefficient of .88.
The Spearman rho between the DSM–IV APD scale and the PCL–R total score was .85 and the Pearson r was .84 (95% CI = .76–.90). It was of interest to correct for the attenuation in the correlations caused by measurement error. Using the correction for attenuation formula, the estimated Pearson correlation between the DSM–IV APD true score (scored as a scale) and the PCL–R true score rose to .96 (95% CI = .94–.98). This correlation between the PCL–R and DSM APD criteria was higher than those reported elsewhere (e.g., Widiger et al., 1994, reported correlations ranging from .48 to .86 using a modification of the PCL–R), although earlier reports did not correct for attenuation. Furthermore, the factor score correlations for the PCL–R followed those expected from earlier studies—Factor 2 was more highly correlated with the DSM APD scale than was Factor 1.

When the diagnosis of APD was rendered as instructed in the DSM (ratings of 2 on at least 3 criteria and the presence of conduct disorder before Age 15), the correlation with PCL–R score was .68 ($p < .0001$), a correlation similar to that reported by other investigators (Hare, 1985; Hare, Hart, & Harpur, 1991; Harris et al., 1994; Widiger et al., 1996). The difference between the present correlations ($r = .68$, for dichotomous APD diagnosis, and $r = .84$, for continuous scoring) was significant ($p < .01$; Meng, Rosenthal, & Rubin, 1992), demonstrating that the association between these two measures can be significantly increased by simply changing the scoring strategy. The intercorrelation of the two PCL–R factors was higher (.76; 95% CI = .63–.84) than usually reported (Hare, 1991), although other investigators, using the 22-item PCL, have reported intercorrelations as high as .71 (Kosson, Kelly, & White, 1997).

### Study 2

Study 1 demonstrated that using the APD criteria as a scale rather than a nonempirical dichotomy resulted in a substantially higher relationship between DSM APD and Hare psychopathy. That is, the results indicated that individuals were ranked in the same order by both assessment approaches. The second study was an extension and expansion of the first. We examined the relation between DSM APD criteria and PCL–R scores and a variable theoretically related to both—violent recidivism. As well, we sought evidence of a taxon underlying APD, using data from the same participants previously used to find evidence of a taxon underlying psychopathy as measured by the PCL–R (Harris et al., 1994).

Many researchers have concluded that lifelong persistent offenders differ from other offenders in fundamental ways, but rarely has the question as to whether the construct of lifelong persistent antisociality should be construed as taxonic or nontaxonic been addressed empirically. Meehl and Golden (1982) described a taxon as an “entity, type, syndrome, species, disease, or more generally, a nonarbitrary class (p. 127).” Some generally accepted taxa are sex and species. Within psychopathology, there is evidence of taxonicity for a growing number of disorders. For example, evidence has been reported for a “schizotypy” taxon underlying schizophrenia (Blanchard, Gangestad, Brown, & Horan, 2000; Golden & Meehl, 1979; Korfine & Lenzenweger, 1995; Lenzenweger & Korfine, 1992; Tyrka et al., 1995), a dementia taxon (Golden, 1982), an endogenous depression taxon (Grove et al., 1987; Haslam & Beck, 1994), a pathological dissociation taxon (Waller, Putnam, & Carlson, 1996; Waller & Ross, 1997), and a taxon underlying bulimia nervosa (Gleaves, Lowe, Snow, Green, & Murphy-Eberenz, 2000).

Although PCL–R scores are quasicontinuous (as are APD scores when the criteria are used as a scale), this proves nothing about the taxonic or nontaxonic nature of the construct. Where there is covariation among indicators at one or more steps in the causal chain between a taxon and the observed behavioral and personality characteristics it influences (plus measurement error), a taxon can yield continuous scores (Meehl & Golden, 1982). For example, height might be an indicator of (i.e., correlated with) sex, but even though height is continuous, sex is taxonic.

The logic behind taxometric techniques is straightforward. It is conjectured, for example in Maxcov-Hitmax (below), both within the taxon and within the complement class, that the indicators of the taxon are relatively uncorrelated with each other. Compared to either class separately, greater covariation results from mixing individuals from the taxon and complement together. The conjecture of zero correlation within taxonic classes is used to obtain within-class parameter estimates (e.g., group means), but the conjecture itself can be tested and modified or discarded on the basis of empirical findings.

To take a simple example, imagine one wished to detect a sex taxon among a group of male and female 6-year-olds, but knew only a few indirectly informative things about each child—hair length, favorite toy, commonly worn clothing, and so on. One would expect that, among a group of girls, that toy choice, hair

---

**Table 1**

*Study 1: Descriptive Statistics and Intercorrelations (Pearson Coefficients) of Main Variables*

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>Alpha</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. DSM–IV APD (scale)</td>
<td>8.49</td>
<td>3.70</td>
<td>.76</td>
<td>.72</td>
<td>.83</td>
<td>.84</td>
<td>.69</td>
</tr>
<tr>
<td>2. PCL–R Factor 1</td>
<td>7.55</td>
<td>4.76</td>
<td>.90</td>
<td>.76</td>
<td>.92</td>
<td>.58</td>
<td></td>
</tr>
<tr>
<td>3. PCL–R Factor 2</td>
<td>8.79</td>
<td>4.85</td>
<td>.82</td>
<td>.94</td>
<td>.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. PCL–R total score</td>
<td>18.41</td>
<td>10.10</td>
<td>.91</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. DSM–IV APD (category)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. DSM–IV = Diagnostic and Statistical Manual of Mental Disorders (4th ed.; American Psychiatric Association, 1994); APD = antisocial personality disorder; PCL–R = Psychopathy Checklist—Revised (Hare, 1991).*
length, and how often the child wears pink would be uncorrelated (or they would have low intercorrelations). In a group of boys, similar low intercorrelations among these indicators would be expected. On the other hand, in a group composed equally of boys and girls, hair length, toy choice, and type of clothing indicators should be much more highly intercorrelated because they are effects of (or are correlated with) a common cause, the sex taxon.

The Maxcov-Hitmax method (Meehl & Golden, 1982; Meehl & Yonce, 1996; Waller & Meehl, 1998) is a taxometric procedure that can be performed with a set of dichotomous items that materially differentiates the taxon from the complement class on average (as long as all are not derived from a scale specifically constructed to be highly unifactorial). Gangestad and Snyder (1985) recommended that eight indicators be used when applying Maxcov-Hitmax to a multi-item scale. This recommendation is based on the rationale that either seven or eight items are probably the minimum number of items that would still give an acceptable number of intervals for graphing. A high number of intervals is desirable but the inclusion of too many item pairs with substantial within-class correlations should be avoided. The items are chosen such that each correlates highly with the total score on the instrument, but they do not have similar manifest content and are not too highly correlated with one another. One can apply Maxcov-Hitmax to multi-item scales by removing two of the eight items (for example) from the scale and constructing a subscale from the remaining items. Then the total sample of individuals is divided into subsamples, by dividing the range of subscale scores into intervals (e.g., one for each possible score on a 0–6 subscale). If a taxon exists, the covariance between the two items set aside, plotted against the sum of the remaining six items, should be peaked, with a maximal value where individuals from both classes are mixed together nearest that 6-item subscale score at which individuals are mixed nearest to a 50:50 ratio, and much lower values at the extremes.

The MAMBAC (mean above minus below a cut) method is a taxometric procedure that can be used with two quasicontinuous variables that putatively discriminate a taxon from its complement class. The MAMBAC procedure requires the evaluation of the mean that participants exhibit on one variable above versus below a series of cut scores on a second variable (Meehl, 1992; Meehl & Yonce, 1994). This mean difference reaches a maximum at that cut point where the false-positive rate equals the false-negative rate, whereas the difference becomes progressively smaller as the cut off moves toward higher or lower scores. Thus, an inverted U-shaped graph indicates a taxon. By contrast, if there is no taxon, and scores on each scale are unimodally distributed along a continuum, the MAMBAC graph will show an upright U-shaped curve.

The goodness-of-fit index (GFI) is one of several consistency tests that allows an evaluation of the degree to which the covariation between indicators generated by the taxonic model fits the observed pattern (Waller & Meehl, 1998). In simulation studies of the GFI, Waller and Meehl (1998) found that high GFI values were produced in taxonic but not nontaxonic samples. The GFI can be interpreted as a multivariate $R^2$; values of this index range from 0.00 to 1.00, with higher values indicating better fit. In taxonic samples with little or no within-class interitem covariance, GFI values were generally greater than .90.

Meehl and colleagues (Meehl, 1995; Meehl & Yonce, 1994, 1996; Waller & Meehl, 1998) have reported strong evidence of the validity of these methods from computer simulations using taxonic and nontaxonic data sets with a variety of base rates and sample sizes. However, Miller (1996) raised concerns about the validity of the Maxcov-Hitmax procedure, arguing that evidence from one taxometric graph is not sufficient evidence of a taxon, because spuriously taxonic-looking graphs can occur. Meehl and colleagues (Meehl, 1996; Meehl & Yonce, 1996; Waller & Meehl, 1998) agreed with this point and argued (consistent with their previous writings) that a taxon should be inferred only when multiple taxometric procedures agree. They also argued that the validity of results should be judged by numerical consistency rather than by conventional statistical significance. If several taxometric methods indicated the presence of a taxon, and parameter estimates arrived at in different ways (derived within and across taxometric procedures) were numerically consistent, then there would be strong grounds for believing the taxon existed (Meehl, 1995; Waller, personal communication, February 7, 2000; Waller & Meehl, 1998), and it would be a “damn strange coincidence” otherwise (Meehl, 1990a, 1990b; Salmon, 1984). Miller (1996) also questioned the suitability of dichotomous indicators in the Maxcov-Hitmax method. Meehl and Yonce (personal communication, February 8, 2000) noted, however, that the weight of the evidence to date supports the use of dichotomous indicators in the Maxcov-Hitmax procedure, when consistency tests are properly used. As well, a recent study (Ruscio, 2000) demonstrated that Maxcov-Hitmax can discern latent taxa when adapted for use with dichotomous indicators.

Study 2 examined (like Study 1) the relationship between DSM–III–R APD scored as a scale and PCL–R score. In addition, taxometric analyses were performed using DSM APD items. These taxometric analyses were similar to those conducted using PCL–R items and the non-PCL–R measure (juvenile antisocial behavior items) in our previous taxometric study (Harris et al., 1994) that gave clear taxonic results. If APD indicators show evidence of taxonicity, similar to that exhibited by PCL–R items (as we previously reported; Harris et al., 1994), this strengthens our conclusion that both measures tap the same underlying taxonic construct. Such findings would suggest that past disagreements about the nature and relationship of these two constructs were in part due to measurement problems, not to any fundamental differences in the constructs.

Method

Participants

Data from two previously reported follow-up studies (Rice et al., 1992; Rice, Harris, Lang, & Bell, 1990) were combined to yield observations on 684 male offenders (also reported in Harris et al., 1994). Of the total, 361 were men treated in a maximum security psychiatric institution between 1965 and 1980. Of these, 297 or 45% of the total, had been found not guilty by reason of insanity, and the remainder had been convicted and then sent to the institution for treatment while under sentence. In Ontario, during the time these participants were identified, more than 40% of the offenders found not guilty by reason of insanity would not have met DSM–III criteria for any psychotic diagnosis (Rice, Harris, Lang, & Bell, 1990). An additional 288 participants had been admitted to the study institution only for short (<60 days) pretrial assessments. Almost all (>95%) of these latter participants were convicted and incarcerated elsewhere, and none returned to the institution for treatment. Each participant in the second group was matched to a participant in the first on several characteristics: age, index
offense, year of index offense, violent criminal history, and nonviolent criminal history (see Rice, Harris, Lang, & Bell, 1990; Rice et al., 1992). The matching was used to evaluate treatment at the study institution. These 684 participants included 184 sex offenders (Rice & Harris, 1997) assessed at the maximum security institution’s sexual behavior laboratory. The present sample was, therefore, not representative of criminals in general and substantially overlapped (> 80%) with the participants used in our earlier report on psychopathy as a taxon (Harris et al., 1994). Data reported in the previous studies (Harris et al., 1994; Rice, Harris, & Quinsey, 1990; Rice et al., 1992) indicated that these participants were a group of serious offenders with many violent offenses and extensive histories of antisocial behavior as children and adults.

Variables

All variables (including PCL–R scores), except those pertaining to recidivism, were coded retrospectively and exclusively from descriptive material contained in institutional files. This coding was completed by a team of three research assistants who had extensive experience and training (mean greater than 3 years) as data coders for similar research projects. The institutional files included information from a variety of sources (e.g., psychosocial histories, information from other institutions, police reports, psychological test reports, questionnaires from patients’ families). Most of the variables used in this study are self-explanatory. However, a few require some additional explanation.

Elementary school maladjustment was rated on a 4-point scale (0 = none, 1 = serious discipline problems, 2 = serious discipline problems and/or attendance problems). Childhood behavior problems was the sum of items endorsed for the 12 problem behaviors noted before Age 15 for a DSM–III diagnosis of antisocial personality disorder. Teen alcohol abuse was rated on a 4-point scale (0 = never drank, 1 = serious drinking problem). Separation from parents was scored as present if it occurred due to divorce, abandonment, or institutionalization (of either parent or child) before the participant turned 16. When individual items on the PCL–R could not be coded (< 5%), scores were prorated. Although the PCL–R is often completed with a clinical interview, scores in this study were, as permitted by the manual (Hare, 1991), based entirely on file information.

To examine variables pertaining to childhood and early adolescence, supplementary coding using the same methods was performed by different research assistants who were blind to the scores on the variables just described. This supplementary coding included the coding of the criteria pertaining to DSM–IV conduct disorder. DSM–IV APD criteria were scored by using PCL–R item scores in the case of the few identical items (e.g., lack of remorse, conning, and manipulative and pathological lying) or separate study variables (e.g., history of criminal conduct as grounds for arrest, rating of aggression as an adult). One APD item (reckless disregard for the safety of self or others) could not be coded from the data available. Adult items as well as conduct disorder items were converted to standard scores that were summed to arrive at a final scaled rating for DSM–IV APD.

*Childhood and Adolescent Taxon Scale (CATS).* As discussed above, a previous study (Harris et al., 1994) indicated that eight individual variables could identify members of the psychopathy class. These variables were: elementary school maladjustment, teen alcohol abuse score, childhood aggression, childhood behavior problems, parental alcohol problem, suspended or expelled from school, separation from parents before Age 16, and arrested under Age 16. In the Harris et al. (1994) study, each variable could identify members of the psychopathy class. These variables were dichotomized and the sum, ranging from 0 to 8, was related to the probability that the participant was a member of the psychopathy class.

A separate team of research assistants gathered recidivism data obtained from the files of the Coroner’s Office, the Ontario Review Board (which maintains information about every insanity acquittee in Ontario), the Royal Canadian Mounted Police (a national database of criminal arrests and convictions), the National Parole Service of Canada, and provincial correctional and parole systems. Childhood history, adult adjustment, offense variables, and assessment variables were coded before recidivism data in order to prevent inadvertent contamination of the historical variables by raters’ knowledge of later crimes.

In coding recidivism, participants were classified as violent failures if they incurred any new charge for a criminal offense against persons or were returned to a maximum security institution for violent behavior against persons who, in the judgment of the first team of raters, would otherwise have resulted in a criminal charge for an offense against persons. Violent acts included assaults, sexual assaults, armed robbery, forcible confinement, threatening, and pointing a firearm, but did not include possession of a weapon, robbery, or arson. A participant was deemed to have had an opportunity to recidivate when he was released to the street or was placed in a halfway house or an open psychiatric ward. The mean length of opportunity to recidivate was 123.0 (SD = 73.4) months.

*Results and Discussion*

*Interrater Reliability and Intercorrelations*

Twenty participants were randomly chosen for an interrater reliability check. One rater (the one who coded the largest number of participants) rated all 20 of the participants selected for the reliability check, and each of the other two raters rated 10. Pearson correlation coefficients were computed for continuous variables (overall correlation = .90) and kappa for categorical variables (overall kappa = .83).

The interrelationships of the main variables are shown in Table 2. As in Study 1, the relationship between PCL–R score and the APD scale was very high (Pearson $r = .85$, Spearman $p = .89$), even without correction for attenuation due to unreliability. As well, even though all the present participants were serious criminal offenders and might be expected to exhibit some restriction of range in scores, the intercorrelations in Table 2 indicate that the...
two measurement approaches again ranked the participants in essentially the same order.

**Taxometric Analyses**

As described in the introduction, a Maxcov-Hitmax analysis was applied to the scale scores for *DSM-IV* APD and the results, averaged over the 28 possible computations, are shown in Figure 1. For comparison purposes, Figure 1 also shows the same analyses applied to the PCL–R and the CATS. As the figure shows, the APD scale produced a result consistent with a taxon, but perhaps not as strongly as that provided by the CATS. As mentioned in the introduction, we also applied the MAMBAC method. The CATS and the CD items were used in conjunction with the APD items. As discussed earlier, an inverted U-shaped MAMBAC curve indicates a taxon, whereas an upright U-shape indicates a nontaxonic construct. The results for the present study are shown in Figure 2. According to the standards set by Monte Carlo simulations provided by Meehl and Yonce (1994), the current MAMBAC analyses gave evidence of an underlying taxon for the APD indicators. Because the CATS had earlier given the clearest evidence of a taxon, the most informative parts of Figure 2 are probably those analyses that use the CATS as one of the two measures—row 3 and column 3 in the figure, and all four of the patterns have the taxonic inverted U shape. Overall, none of the analyses using the scaled APD were consistent with a nontaxonic entity. Consistency tests performed with the APD and CATS measures also bolstered the position that an underlying taxon exists (GFI = .93; Waller & Meehl, 1998), especially given the nuisance covariance that existed between these two measures.

There is no agreed-upon (e.g., mathematically derived) method for testing whether a plot produced by these analyses was taxonic (Miller, 1996). Therefore, we asked eight nonresearchers from various disciplines, with no knowledge of taxometric analyses, to compare these graphs with sets of simulated taxonic and nontaxonic plots (with the same gross distributional characteristics as our own data), provided by Meehl and Yonce (1994, 1996). Participants were asked to decide, for each plot, which set it most resembled. All eight raters placed all of our MAMBAC plots with the taxonic plots provided. The probability of this degree of agreement occurring by chance is less than .001.

Because indicators from both statistical estimation and testing procedures gave evidence of an underlying taxon, it was also of interest to examine the relationship between the two measures when both were used to empirically categorize participants as taxon or complement members. Using the taxometric methods described by Golden (1982), Bayesian estimates were used to classify participants as members of the taxon by each of the two assessment tools (APD criteria or PCL–R). The agreement between the two methods of who was and who was not a taxon member was very high (κ = .82), and when corrected for attenuation reached unity (see Grove, 2000, for the derivation of a disattenuation formula for kappa when used to relate two hypothetical latent taxa).

**Subsidiary Analyses**

In assessing the risk and supervision needs of offenders, criminal justice authorities face resource limitations. Thus, we tried to determine what minimal set of variables might best detect the single discrete entity that appeared to underlie the primary measures in this study. To evaluate potential indicators, we tested each of the 20 PCL–R items, all eight CATS variables, the 15 *DSM-IV* conduct disorder items, and six of the seven *DSM-IV* APD items (as mentioned, we did not have “reckless disregard for safety of self or others” item). The dependent measure was a single dichot-

---

**Table 2**

<table>
<thead>
<tr>
<th>Measure</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Violent recidivism</td>
<td>.25</td>
<td>.39</td>
<td>.36</td>
<td>.32</td>
<td>.37</td>
<td>.32</td>
</tr>
<tr>
<td>2. PCL–R Factor 1</td>
<td>.59</td>
<td>.85</td>
<td>.37</td>
<td>.76</td>
<td>.35</td>
<td></td>
</tr>
<tr>
<td>3. PCL–R Factor 2</td>
<td>.92</td>
<td>.51</td>
<td>.77</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. PCL–R total score</td>
<td>.49</td>
<td>.85</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Conduct disorder (scale)</td>
<td>.81</td>
<td>.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. <em>DSM-IV</em> APD (scale)</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Child and Adolescent Taxon Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Listwise deletion for missing data (*N* = 436). PCL–R = Psychopathy Checklist—Revised (Hare, 1991); *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; APA, 1994); APD = antisocial personality disorder.

---

7 When the base rate of taxon membership is close to .5, the inverted U has a horseshoe appearance. As the base rate approaches .25, the inverted Us much more resemble Js.
omous variable—whether the participant was identified as having a very high probability of being a taxon member (> .99 using Bayesian estimates) by both the APD and the PCL–R items or not. We conducted a forward binary logistic regression (α to enter or remove = .01) which identified 12 items achieving high classification accuracy (98.4% correct classification), $\chi^2(12, N = 329) = 508$, $p < .0001$, $R^2 = .95$. The variables selected were elementary school maladjustment score (from the CATS), the score representing the frequency and severity of adult violence (from DSM–IV APD), impulsivity, lying, conning and manipulation, lack of remorse, callousness, and irresponsibility (from the PCL–R), early behavior problems and shallow affect (from the PCL–R), and initiating fights and using a weapon in fights (from DSM–IV conduct disorder).

We also performed a similar analysis (α to enter and remove = .05) with violent recidivism (scored dichotomously) as the dependent variable. Interestingly, the results suggested that a very small set of items could have predicted violent recidivism just as well as any other study variable, such as the full PCL–R, for example (see Table 2). The analysis yielded a multiple $R = .33$, $p < .0001$, 67% correct classification, $\chi^2(7, N = 306) = 126$, $p < .0001$, $R^2 = .20$. Most interesting was the result that the analysis selected items from each of the CATS (elementary school maladjustment, separation from parents), DSM–IV APD criteria (the extent and severity of criminal history), DSM–IV conduct disorder (running away from home), and the PCL–R (callousness, parasitic lifestyle, and sexual promiscuity). The bivariate correlations for the variables from these regression analyses are shown in Table 3. It is also of interest to note that PCL–R Factor 2, PCL–R total score, and DSM–IV APD scale score each were just as highly correlated with violent recidivism as were combinations of variables (see Table 2).

In addition to these results, some other relations were of interest. In predicting dichotomous violent recidivism, a forward logistic regression procedure (alpha to enter and remove = .05) selected only PCL–R Factor 2 and conduct disorder scores (69% correct classification), $\chi^2(2, N = 431) = 78.1$, $p < .0001$, $R^2 = .22$. Conduct disorder and both PCL–R factors accounted for essentially all the outcome variance in predicting DSM–IV APD score (multiple $R = .97$). Of course, one would expect shrinkage in the $R^2$s from these regression equations upon cross-validation and establishing the validity of the final equations from these analyses awaits another study with a new sample of offenders (see Lees & Neufeld, 1994).

General Discussion and Conclusions

The results from two samples of male offenders indicated that when the PCL–R (whether scored from file alone or from file and interview) and DSM APD criteria are each scored as continuous measures the association between them is extremely high. Each method ranked participants in essentially the same order, with any differences attributable almost entirely to measurement error. This high rank-order correlation means that, assuming the use of appropriate cut-off scores, individuals would be categorized nearly

---

8 A small number of participants (< 10%) were classified as taxon members by one method and as members of the complement by the other. We performed the analysis treating these participants as complement members and also eliminating them from the analysis and obtained substantially the same results.

9 Nagelkerke $R^2$.

10 Although these items were all selected from the PCL–R, note that all are essentially equivalent to corresponding items in DSM–IV APD criteria.
observed between the PCL–R and APD, exactly the same high degree of association was indicated by the two approaches. Changing the diagnosis of psychopathy and APD from nonempirically derived categories to continuous measures increased their interrelationship substantially and significantly. The results mean that persistently antisocial offenders as indexed by both PCL–R and APD criteria are of greater importance than the more behavioral characteristics for taxon identification or behavioral predictions.

How is it possible to reconcile the conclusions reached by Cooke and colleagues with the results of the present study? The contradiction is, in our view, more apparent than real. Item response theory, as applied by Cooke and colleagues, measures the manner and degree to which particular items locate individuals on a latent continuum. They do not test the predictive validity of the items, individually or collectively. Thus, interpersonal and affective PCL–R items might ably tap something fundamental but less
strongly related to external criteria than the more behavioral items.
In fact, PCL–R Factor 2 has commonly been found to be a better predictor of general, violent, and sexual recidivism (Quinsey, Harris, Rice, & Cormier, 1998; Salekin et al., 1996; Serin et al., 1990). Although it is not yet certain what traits or characteristics will prove optimal in identifying the discrete entity evident in the present results, items relating to early, persistent antisocial behavior appear to work best for the assessment of violence risk. It is possible that for other purposes, such as assessing treatment appropriateness, the affective or interpersonal items of the PCL–R will prove more useful than behavioral indicators prominent in DSM–IV.

These considerations provide an important part of the explanation of the ability of such actuarial instruments as the Violence Risk Appraisal Guide (VRAG; Harris et al., 1993; Quinsey et al., 1998) to predict violent recidivism among violent adult offenders. Because the VRAG contains PCL–R and other items that measure juvenile antisocial and aggressive behavior, offenders who score high on such scales were, by definition, antisocial and aggressive children whose antisocial and aggressive behaviors persisted into adulthood. This interpretation of the success of actuarial instruments that contain items derived from childhood predictors raises theoretically and practically important questions. Can a persistent antisociality or conduct disorder taxon be identified in childhood? If so, what proportion of taxon members later desist from aggressive and antisocial behavior? And what characteristics do these desisting taxon members exhibit in early childhood and late adolescence? Researchers are attempting to identify the “fledgling psychopath” (Frick, O’Brien, Wootton, & McBurnett, 1994; Lynam, 1996) and have demonstrated that an antisociality taxon can be found in a community sample of school-aged boys (Skilling, Quinsey, & Craig, 2001). Finding a discrete entity underlying antisociality in a sample of children addresses the generalizability of these taxon findings and strengthens the suggestion that persistent antisociality is a distinct life history strategy beginning in early childhood and often resulting in a lifetime of criminal and violent behavior (Harris, Rice, & Lalumière, 2001; Harris, Skilling & Rice, 2001; Lalumière, Harris, & Rice, 2001; Quinsey, 1995; Quinsey et al., 1998).

References


Korfine, L., & Lenzenweger, M. F. (1995). The taxonicity of schizotypy:


Harup, T. J., Hakstian, R., & Hare, R. D. (1988). Factor structure of the


Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of

taxon: Evidence that psychopaths are a discrete class. Journal of Con-
sulting and Clinical Psychology, 62, 387–397.


psychopaths following conditional release from prison. Journal of Con-
sulting and Clinical Psychology, 56, 227–232.


Korfine, L., & Lenzenweger, M. F. (1995). The taxonicity of schizotypy:

construct validity of psychopathy in Black and White male inmates:
Three preliminary studies. Journal of Abnormal Psychology, 99, 250–
259.

Lalumière, M., Harris, G. T., & Rice, M. E. (2001). Psychopathy and

inference to the limits of quantitative methods: A formal appeal to
practice what we consistently preach. Canadian Psychology, 35, 268–
282.


Loober, R. (1982). The stability of antisocial and delinquent child behavior:


Lynam, D. R. (1996). Early identification of chronic offenders: Who is the

Meehl, P. E. (1990a). Appraising and amending theories: The strategy of


Meehl, P. E. (1992). Factors and taxa, traits and types, differences of


gist, 51, 1184–1186.

Kendall & J. N. Butcher (Eds.), Handbook of research methods in

taxonomy with two quantitative indicators using means above and below
a sliding cut (MAMBC procedure). Psychological Reports, 74, 1089–
1274.

taxonomy using covariance of two quantitative indicators in successive
intervals of a third indicator (MAXCOV procedure). Psychological
Reports, 78, 1091–1227.


Miller, M. B. (1996). Limitations of Meehl’s MAXCOV-HITMAX pro-

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent anti-


security therapeutic community for psychopaths and other mentally
disordered offenders. Law and Human Behavior, 15, 625–637.

Rice, M. E., Harris, G. T., Lang, C., & Bell, V. (1990). Recidivism among
male insanity acquittees. Journal of Psychiatry and the Law, 18, 379–
403.

assessed in a maximum security psychiatric facility. Journal of Inter-

Robins, L. N. (1978). Etiological implications in childhood histories relating
to antisocial personality. In R. D. Hare & D. Schalling (Eds.), Psychopathic behavior: Approaches to research (pp. 255–272). Chi-
chester, United Kingdom: Wiley.

analysis of antisocial personality disorder: DSM–IV and beyond. Law
and Human Behavior, 18, 471–484.

Ruscio, J. (2000). Taxometric analysis with dichotomous indicators: The

Gender differences in the antisocial personality disorder criteria to
Psychopathy Checklist—Revised scores. Journal of Personality Disor-
ders, 12, 69–76.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Check-