Using the PCL-R to Help Estimate the Validity of Two Self-Report Measures of Psychopathy With Offenders

Article *in* Assessment · November 2009

DOI: 10.1177/1073191109351715 · Source: PubMed

CITATIONS
138

READS
1,045

7 authors, including:

Jennifer L Skeem
University of California, Berkeley
118 PUBLICATIONS **9,581** CITATIONS
SEE PROFILE

John F Edens
Texas A&M University
194 PUBLICATIONS **9,187** CITATIONS
SEE PROFILE

Monica Epstein
University of South Florida
20 PUBLICATIONS **1,099** CITATIONS
SEE PROFILE

Christopher J Patrick
Florida State University
279 PUBLICATIONS **16,351** CITATIONS
SEE PROFILE

Using the PCL-R to Help Estimate the Validity of Two Self-Report Measures of Psychopathy With Offenders

Norman G. Poythress, Scott O. Lilienfeld, Jennifer L. Skeem, Kevin S. Douglas, John F. Edens, Monica Epstein and Christopher J. Patrick

Assessment 2010 17: 206 originally published online 14 November 2009
DOI: 10.1177/1073191109351715

Using the PCL-R to Help Estimate the Validity of Two Self-Report Measures of Psychopathy With Offenders

Assessment
17(2) 206–219

DOI: 10.1177/1073191109351715
http://asmnt.sagepub.com
SAGE

Norman G. Poythress,[1] Scott O. Lilienfeld,[2] Jennifer L. Skeem,[3] Kevin S. Douglas,[4] John F. Edens,[5] Monica Epstein,[1] and Christopher J. Patrick[6]

Abstract

Two self-report measures of psychopathy, Levenson's Primary and Secondary Psychopathy scales (LPSP) and the Psychopathic Personality Inventory (PPI), were administered to a large sample of 1,603 offenders. The most widely researched measure of criminal psychopathy, the Hare Psychopathy Checklist–Revised (PCL-R), served as a provisional referent for estimating the construct validity of these self-report measures with offenders. Compared with the LPSP, the PPI displayed higher zero-order correlations with the PCL-R, better convergent and discriminant validity, and more consistent incremental utility in predicting PCL-R scores. Furthermore, using a variant of Westen and Rosenthal's approach to evaluating the construct validity of a new measure, compared with the LPSP, the PPI's pattern of associations with measures of 35 external criterion variables was more similar to the pattern observed for the PCL-R. Results generally provide stronger support for the validity of the PPI than the LPSP in offender populations using the PCL-R as a provisional benchmark, particularly for assessing interpersonal and affective features of psychopathy.

Keywords

self-report psychopathy, Psychopathic Personality Inventory, Levenson Primary and Secondary Psychopathy scales

Assessing Psychopathy: The Hare Psychopathy Checklist

For more than a quarter century, the Hare Psychopathy Checklist (PCL; Hare, 1980) and the Hare Psychopathy Checklist–Revised (PCL-R; Hare, 1991, 2003) have been the measures of choice for assessing psychopathic features in offender and forensic psychiatric samples. The PCL-R is a 20-item clinician-rated measure scored on the basis of a semistructured interview that queries participants about many of the core interpersonal (e.g., superficial charm, conning and manipulative behavior) and affective (e.g., callousness, lack of empathy, lack of remorse or guilt) features of psychopathy, as well as the deviant lifestyle (e.g., irresponsible behavior, parasitic lifestyle) and criminal behaviors often associated with this condition. The PCL-R also requires that interviewers consider corroborative data (e.g., file information) when rating items. A large body of literature attests to the validity of the PCL-R. For example, PCL-R scores are robustly associated with personality measures of poor impulse control and reliably predict responses on laboratory tasks such as measures of passive avoidance learning (Newman & Kosson, 1986) and emotional responding to startle probes (Patrick, 1994, 2001). The PCL-R is popular with forensic clinicians who conduct risk assessments (Lally, 2003) because of its utility in predicting violent and criminal recidivism (Walters, 2003).

However, just as any measure, the PCL-R has limitations. Pragmatic limitations include that it is time-consuming (a typical administration often requires 90 minutes or more), cannot be administered in groups, requires extensive training to administer and score, and is of questionable validity in settings in which collateral data (e.g., institutional file records) are unavailable or of poor quality. Furthermore,

[1]University of South Florida, Tampa, FL, USA
[2]Emory University, Atlanta, GA, USA
[3]University of California at Irvine, Irvine, CA, USA
[4]Simon Fraser University, Burnaby, British Columbia, Canada
[5]Texas A&M University, College Station, TX, USA
[6]University of Minnesota, Minneapolis, MN, USA

Corresponding Author:
Norman G. Poythress, Department of Mental Health Law and Policy, Florida Mental Health Institute, University of South Florida, Tampa, FL 33612, USA
Email: poythres@fmhi.usf.edu

because it was developed for use with offenders and scoring criteria for several items heavily reference criminal behavior, it is not suitable for use with noncriminal populations.

Others have expressed concerns that the PCL-R conceptualization of psychopathy is both under- and overinclusive. For example, the PCL-R does not assess for (the absence of) interpersonal anxiety or neurotic habits, a hallmark feature of psychopathy according to Cleckley (1941/1964) and considered a critical marker for distinguishing between primary and secondary subtypes of psychopathy (Karpman, 1949). As a result, some investigators have found it necessary to supplement PCL-R assessments with scores from a measure of trait anxiety (e.g., Newman & Schmitt, 1998; Skeem, Johansson, Andershed, Kerr, & Louden, 2007). Conversely, because psychopathy is not thought to be limited to offenders, others have questioned whether the PCL-R's explicitly criminological items are appropriate indicators of the construct (e.g., Andrade, 2008; Blackburn, 2007).

Alternative Measures of Psychopathic Features

Several researchers have developed self-report measures of psychopathy that could serve as alternatives to the PCL-R. Self-report measures are economical, easily administered, and can include scales for the detection of response sets and styles (e.g., positive impression management, malingering). Furthermore, compared with interviews, self-report measures may permit a more comprehensive and sensitive assessment of subjective emotional dispositions, such as guiltlessness and fearlessness (Lilienfeld & Fowler, 2006). However, self-report measures also have potential disadvantages. For example, because psychopathic individuals are prone to deception and may lack insight into the nature and extent of their deficits, the ability of self-report measures to detect psychopathy may sometimes be compromised (Lilienfeld, 1994). It may also be paradoxical to ask individuals to report on the frequency of emotions (e.g., guilt) that they have rarely, if ever, experienced and perhaps do not comprehend psychologically.

In this study, we examine the construct validity of two self-report scales that have shown promise for assessing psychopathic traits, the Levenson Primary and Secondary Psychopathy scales (LPSP; Levenson, Kiehl, & Fitzpatrick, 1995) and the Psychopathic Personality Inventory (Lilienfeld & Andrews, 1996). Both were designed to assess psychopathic features in noninstitutional samples; thus, both measures focus on broad personality traits and dispositions and exclude explicitly criminological items. Somewhat different strategies were employed in developing these two measures. The LPSP is tied closely to the PCL-R operationalization of psychopathy, as its items "were designed to produce by the means of a self-report procedure, two factors similar to those produced by the Hare Psychopathy Checklist" (Levenson et al., 1995, p. 152). In an undergraduate sample, a principal components analysis yielded two dimensions—the Primary psychopathy scale, intended to assess features captured by PCL-R Factor 1 (e.g., lack of guilt, callousness) and the Secondary psychopathy scale, presumably analogous to PCL-R Factor 2 and intended to capture features of neurotic turmoil coexisting with impulsive behavior.

In contrast, the PPI was developed in reference to the broader psychopathy literature, and its items were written to assess more than 30 focal constructs potentially relevant to this condition (e.g., lack of guilt, disloyalty, poor impulse control, risk taking, externalization of blame). Successive factor analyses of the PPI item pool across three independent undergraduate samples revealed eight lower order dimensions ostensibly assessing differing facets of psychopathy. Subsequent analysis of these eight scales in a community sample (Benning, Patrick, Hicks, Blonigen, & Krueger, 2003) identified two higher order factors labeled Fearless Dominance and Impulsive Antisociality, which are roughly analogous to the PCL-R Factors 1 and 2 (but see Neumann, Malterer, & Newman, 2008 for an alternative model of the PPI factor structure).

Although the LPSP and PPI have been used in numerous studies with undergraduate and community samples (see Lilienfeld & Fowler, 2006, for a review), investigators have also begun to explore the utility of these measures with offender samples (e.g., Brinkley, Diamond, Magaletta, & Heigel, 2008; Brinkley, Schmitt, Smith, & Newman, 2001; Poythress, Edens, & Lilienfeld, 1998; Sandoval, Hancock, Poythress, Edens, & Lilienfeld, 2000). Because the PCL-R explicitly was developed to assess psychopathy in offenders and has been extensively studied and validated in this population, it may be viewed as a provisional "benchmark" for comparing the performance of the LPSP and PPI. The more these self-report measures perform like the PCL-R, the greater the likelihood that investigators and clinicians may approach them as efficient alternatives to psychopathy assessment with offenders. To date, however, only a few studies (for the LPSP, Brinkley et al., 2001; for the PPI, Berardino, Meloy, Sherman, & Jacobs, 2005; Edens, Poythress, Lilienfeld, & Patrick, 2008; Malterer, Lilienfeld, Neumann, & Newman, in press) have examined the convergent and discriminant validity of these self-report measures in relation to the PCL-R, and no study has examined their incremental utility above and beyond each other for predicting PCL-R scores. Nor are there data evaluating the extent to which these self-report measures exhibit patterns of associations with external criterion variables similar to that observed for the PCL-R.

With these gaps in the literature in mind, our goals in this study are threefold. First, we first examine the correlations among the PPI, LPSP, and PCL-R both in terms of total and scale scores. Second, using the PCL-R as a provisional

benchmark for the assessment of psychopathy with offenders, we examine the incremental utility of the PPI above and beyond the LPSP, and vice versa, for predicting PCL-R total and factor scores. Third, we adapt the quantitative metrics for construct validation developed by Westen and Rosenthal (2003) to examine whether the PPI or the LPSP more closely approximates the PCL-R in its pattern of correlations with external correlates. As we discuss below, the Westen and Rosenthal framework can be used to compare how well a "new" measure (e.g., the LPSP or PPI) replicates the pattern of statistical associations with criterion measures obtained using a better established measure (e.g., the PCL-R).

For these analyses, we employ the two-factor structures that have been most widely used in research with the PCL-R (Hare, 1991), the PPI (Benning et al., 2003), and the LPSP (Levenson et al., 1995) and which maximize the comparability of these measures.[1] Although not isomorphic across measures, these factors assess the core interpersonal and affective features of psychopathy (Factor 1) and the socially deviant lifestyle with which it is often associated (Factor 2). Our criterion measures represent a broad nomological network of constructs (Cronbach & Meehl, 1955) that bear clinical, conceptual, or theoretical relevance to different conceptualizations of psychopathy.

Method

Participants

Participants included offenders who were court-ordered to community residential drug treatment programs or serving prison sentences in Oregon, Utah, Nevada, and Florida. In addition, we recruited participants from a residential drug treatment program (located within a prison) in Texas. Given our interest in psychopathy, which is more prevalent in men than women (Cale & Lilienfeld, 2002), our recruitment strategy favored the enrollment of male (80%) versus female (20%) participants.

As other investigators have done (e.g., Swogger & Kosson, 2007; Vassileva, Kosson, Abramowitz, & Conrod, 2005), we excluded participants who resided on a designated mental health unit within a prison or who were receiving psychotropic medications for active symptoms of psychosis to minimize the possible impact of symptoms of major mental disorders on protocol responses. Eligibility criteria included the ability to speak English (could be a second language) and having an estimated IQ ≥70 on a screening measure (Quick Test; Ammons & Ammons, 1962) that was administered immediately on enrollment into the study. Individuals in residential drug treatment programs also must have completed any detoxification procedures.

A total of 1,741 participants enrolled, including 1,413 males (81.2%) and 299 females (17.2%), with missing gender data on 29 participants (1.7%). There were 1,079 Caucasians (62%) and 595 African Americans (34.2%), with missing race data for 67 participants (3.8%). In all, 911 participants were recruited from prisons (52.3%) and 830 from drug treatment programs (47.7%). Data were excluded for 39 participants whose *T*-score on either the Inconsistency or Infrequency scale of Morey's (1991) Personality Assessment Inventory (PAI) suggested invalid responding (i.e., $T > 79$, Edens & Ruiz, 2005) and for 41 other participants for whom these profile validity indices were not available (i.e., they did not complete the PAI). We also excluded data from 6 enrollees who failed the IQ screen and 52 enrollees who did not complete the protocol or had extensive missing data. Thus, the present analyses were based on 1,603 participants.[2]

Measures

Measures of Psychopathy

PCL-R (Hare, 1991). The basic features of the PCL-R were described earlier. The PCL-R manual (Hare, 2003) provides a comprehensive review of studies of the measure's psychometric properties and relations with external criteria. For the present study we obtained scores for the two-factor model (Hare, 1991), as this structure best facilitated direct comparisons with the LPSP and PPI. Descriptive and reliability (Cronbach's alpha) data for the PCL-R in the present sample were as follows: Total Score, $M = 22.54$, $SD = 7.49$, $\alpha = .82$; Factor 1, $M = 8.13$, $SD = 4.11$, $\alpha = .81$; Factor 2, $M = 11.19$, $SD = 3.58$, $\alpha = .68$.

High interrater reliability has been reported in previous studies (Hare, 2003); in the present study, on the basis of 51 cases interrater reliability for PCL-R total scores was intraclass correlation coefficient (ICC_1) = .88. To minimize rater drift, the project coordinator (Kevin S. Douglas) conducted regular site visits over the course of the study to observe PCL-R interviews and independently score the PCL-R. These site visits were scheduled to occur approximately every 6 months. It was during these visits that the interrater reliability ratings were made. The project coordinator read the files of, and then observed the interview of, study participants who had been selected in the study's customary (random) manner. Typically, two participants were observed within a 1-day visit (one in the morning, one in the afternoon) at each site. All PCL-R scoring by the project coordinator and the site research assistant (RA) was completed independently of one another, and independent of the study's criterion measures. Spacing site/reliability visits every 6 months ensured that observation and reliability checks were spaced roughly evenly throughout the data

collection period. Thus, the project coordinator rated all 51 interrater reliability cases, each of which also was rated by the different site RAs. In total, there were 14 different RAs who contributed cases both to the overall study, as well as to the interrater reliability procedure. Because there was some RA turnover throughout the study, there were 14 different RAs, who provided between two and eight interrater reliability cases.

LPSP (Levenson et al., 1995). The LPSP is 26-item self-report measure that includes a 16-item Primary scale intended to capture features assessed by PCL-R Factor 1, and a 10-item Secondary scale intended to capture features assessed by PCL-R Factor 2. Respondents rate each item using a 4-point Likert-type scale from 1 (*disagree strongly*) to 4 (*agree strongly*). Descriptive and reliability (Cronbach's alpha) data for the LPSP in the present sample were as follows: Total Score, $M = 55.84$, $SD = 11.69$, $\alpha = .86$; Primary scale, $M = 32.83$, $SD = 8.13$, $\alpha = .84$; Secondary scale, $M = 23.00$, $SD = 5.29$, $\alpha = .73$.

PPI (Lilienfeld & Andrews, 1996). The PPI is a 187-item self-report measure that is scored on a 4-point Likert–type scale. It includes eight clinical scales, seven of which load on two higher order factors labeled Fearless Dominance (FD) and Impulsive Antisociality (IA). In the present sample, descriptive and reliability information for the PPI Total score were as follows: $M = 386.17$, $SD = 41.48$, $\alpha = .91$. For the clinical scales that load on FD (Fearlessness, Social Potency, Stress Immunity), alphas ranged from .80 to .86. For the clinical scales that load on IA (Machiavellian Egocentricity, Carefree Nonplanfulness, Impulsive Nonconformity, Blame Externalization), alphas ranged from .73 to .89.

Measures of Criterion Variables

To compare the self-report psychopathy measures with the PCL-R in their associations with external criteria, we used measures of general personality traits and clinical symptoms, scores on a behavioral task assessing passive avoidance learning, and predictive validity indices related to criminal recidivism on release to the community. The IQ screen was also used as an external criterion measure given that PCL-R Factor 2 scores, in contrast to Factor 1 scores, have typically displayed significant (negative) associations with measures of global intelligence (Harpur, Hare, & Hakstian, 1989).

PAI (Morey, 1991). The PAI is a 344-item, self-report inventory of broadband personality and psychopathology that includes 11 clinical scales (Somatic Complaints, Anxiety, Anxiety-Related Disorders, Depression, Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Alcohol Problems, Drug Problems: alphas in the present sample range from .79 to .94), 5 treatment scales (Aggression, Suicidal Ideation, Stress, Nonsupport, Treatment Rejection: alphas range from .69 to .91), and 2 scales that assess interpersonal style (Dominance, Warmth: both αs = .78). The PAI scales have been widely used in validation studies of psychopathy measures, including the PCL-R (e.g., Douglas, Guy, Edens, Boer, & Hamilton, 2007; Edens, Hart, Johnson, Johnson, & Olver, 2000) and various self-report measures (see e.g., Benning, Patrick, Salekin, & Leistico, 2005; Patrick, Edens, Poythress, Lilienfeld, & Benning, 2006). Also, the PAI's Negative Impression Management (answering so as to yield an unfavorable impression) and Positive Impression Management (answering so as to make a favorable impression) scales are relevant to concerns regarding psychopaths' propensity for deceptive responding; thus these scales were included as criterion measures.

Narrowband measures. In addition to the PAI, we used a number of measures of specific constructs relevant to psychopathy. These were the following.

1. Harmavoidance (HA) scale from Tellegen's Multidimensional Personality Questionnaire (Tellegen, 1982): A reverse measure of the fearlessness construct relevant to Lykken's (1995) theory of primary psychopathy ($\alpha = .86$ in the present sample).
2. Carver and White's (1994) Behavioral Inhibition System (BIS) and Behavioral Activation System (BAS) scales: These scales were designed to measure constructs from Gray's (1987) reinforcement sensitivity theory that have been linked to primary (low BIS) and secondary (high BAS) psychopathy in the theories of Fowles (1980) and Lykken (1995; reliabilities for BIS, BAS Reward, BAS Drive, and BAS Fun Seeking scales were $\alpha = .75$, .82, .85, and .78, respectively).
3. Barratt Impulsivity Scale (Version 11; BIS-11, Barratt, 1994): Karpman (1948) identified impulsivity as a trait potentially useful in distinguishing primary (low impulsivity) from secondary (highly impulsive) psychopathic individuals. Internal consistency of the BIS-11 in the present sample was $\alpha = .86$.
4. Child Abuse and Trauma Scale (CATS; Sanders & Giolas, 1991): Porter (1996) hypothesized that a variant of secondary psychopathy may result from early traumatic experiences (i.e., abuse or abandonment). Internal consistency for the CATS Total score in the present sample was $\alpha = .95$.
5. Dissociative Experiences Scale–Version II (DES-II; Carlson & Putnam, 1993): Porter (1996) argued that the variant of secondary psychopathy that results from early abuse or abandonment should be considered a dissociative condition. The DES-II was administered to assess dissociative features; in the present sample, $\alpha = .93$.

6. Measures of Antisocial Personality Disorder (ASPD): Antisocial behavior is prominent in offender samples and moderately associated with measures of psychopathy, particularly PCL-R Factor 2 (Hare, 1991). We administered two measures of ASPD features, the ASPD module from the Structured Clinical Interview for *DSM-IV* Axis II Personality Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1996) and a self-report measure, the Personality Diagnostic Questionnaire–4 (PDQ-4) ASPD Scale (Hyler, 1994). In the present sample, internal consistency was $\alpha = .83$ for the SCID module and .85 for the PDQ-4 ASPD scale.

Passive Avoidance Learning

Cleckley (1941/1964) identified difficulty learning from punishment as a feature of psychopathic individuals. To assess the relationship between psychopathy as assessed by the PCL-R, LPSP, PPI, and problems with passive avoidance learning, we administered the GoNoGo Task (GNG; Newman & Kosson, 1986) using a laptop computer. On this task, participants complete a block of 40 learning trials during which they learn to discriminate which four (of eight) 2-digit numbers presented on a computer monitor are associated with reward (earning $0.10) and which are associated with punishment (loss of $0.10) in response to a button being pressed on presentation of the stimulus. The key dependent measure is the number of errors of commission (pressing a button in response to a punished number) during a second block of 40 trials. These data were available for 1,302 participants.

Criminal Recidivism

For all participants recruited from drug treatment programs and those released from prison during the course of the study ($n = 1{,}177$), we obtained postrelease arrest records from the National Crime Information Center. These records are compiled on the basis of offenses reported by divisions of law enforcement from all 50 states to the U.S. Federal Bureau of Investigation. The number and type of offenses for each individual were retrieved. We defined violent offenses to include any arrest for murder, manslaughter, assault, robbery, and rape or other sexual assaults. We calculated two dichotomous (Yes/No) recidivism indices for a 1-year follow-up period—whether the person had been arrested for any offense and whether he or she had been arrested for a violent offense.

Procedure

Prior to data collection RAs received extensive training on the entire protocol, including 2.5 days of face-to-face didactic and clinical training from an expert on the PCL-R (Stephen Hart) and subsequent supervised scoring of 10 training tapes obtained from Robert D. Hare. All RAs were required to obtain an $ICC_1 \geq .80$ before starting data collection. Training on the administration and scoring of the SCID-II ASPD module was conducted by Scott Lilienfeld.

At each research site, participants were randomly recruited from lists of individuals who met basic inclusion criteria (i.e., race, English fluency). Enrollment interviews were conducted in a private room, and informed consent was obtained using procedures approved by a university institutional review board. After informed consent was obtained, the IQ screening test was administered.

Participants who had either completed the 10th grade in regular curriculum classes or obtained a GED, and who could read fluently the first few items of the PAI, were allowed to complete the self-report measures alone. Those not meeting these criteria were tested for reading comprehension (Johns, 1997). Self-report items were read aloud to 44 participants.[3] The PAI was administered as a paper-and-pencil measure; the remaining self-report measures were entered into a software program and participants completed these items using a laptop computer. The passive avoidance learning (GNG) task was also administered via laptop computer. The protocol took, on average, 4.5 hours to complete and was typically administered in two sessions. Except at one agency that did not permit participant payments, participants were paid $20.

Analyses

Zero-order correlations were used to evaluate the convergent and discriminant validity among the Total and factor scores of the PCL-R, LPSP, and PPI. Hierarchical multiple regression was used to evaluate the incremental utility of the LPSP and PPI above and beyond one another for predicting PCL-R scores. Statistics developed by Westen and Rosenthal (2003; $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$) were used to investigate the similarity of the pattern of associations with external correlates for the LPSP and the PPI, compared with that of the PCL-R. We summarize the approach derived from Westen and Rosenthal briefly below.

An important indicator of a measure's construct validity is the congruence between its obtained pattern of correlations with measures of other constructs and the pattern predicted by theory (Cronbach & Meehl, 1955; Messick, 1980). Positive correlations with measures of other constructs theoretically associated with the target construct provide evidence of convergent validity. Lack of association with measures of constructs not theoretically associated with the target construct provides evidence of discriminant validity. Thus, researchers commonly present correlations between the measure to be validated with a variety of criterion measures for which a priori associations can be specified. Generally

speaking, the closer the congruence between the patterns of obtained and predicted correlations, the stronger the evidence for construct validity (but see Borsboom, Mellenbergh, & Heerden, 2004, for a competing view).

Although this approach is intuitively straightforward, it may be difficult in practice to obtain a consensus judgment regarding the extent of support for a measure's construct validity when the obtained correlations conform to expectations in varying degrees. Westen and Rosenthal (2003) provided two quantitative metrics for summarizing the pattern of findings in a convergent–discriminant validity array. These metrics, $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$, reflect an extension of contrast analysis more familiarly used in analysis of variance to test a specific hypothesis regarding differences among group means. Both indices require the investigator to specify in advance a set of expected correlations between the target measure and an array of criterion measures. The predicted correlations are converted to lambdas (λs) by subtracting out the mean predicted correlation from each individual predicted correlation, resulting in a set of contrast weights that sum to zero. The obtained correlations are transformed using the Fisher *Zr* transformation. The $r_{\text{alerting-CV}}$ index is obtained by correlating the λ and *Zr* values and indexes the extent to which there is consistency in the ordering of predicted versus obtained correlations.

Although $r_{\text{alerting-CV}}$ is interpretable as an effect size (as is any other correlation), it is characterized as "a rough, readily interpretable index that can alert the researcher to possible trends of interest" (Westen & Rosenthal, 2003, p. 610). The $r_{\text{contrast-CV}}$ index is a more rigorous test of congruence between expected and obtained associations. It is derived from a series of complex calculations (see Westen & Rosenthal, 2003, pp. 617-618) that take into account the median intercorrelations among the criterion measures, the magnitudes of the correlations between the target measure and criterion measures, and sample size.

In the present study, we made two adaptations of $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ in evaluating the validity of the PPI and LPSP. The first adaptation was conceptual. We specified the target pattern of predicted associations on the basis of the observed associations between the PCL-R scales and a broad variety of criterion variables rather than on theoretical grounds. As noted earlier, the PCL-R is widely regarded as the best-validated measure of psychopathy with offender populations and can serve as a provisional referent against which to compare new measures that have been less examined with this group. The observed associations surrounding the PCL-R provide a target array of criterion correlations against which to compare the observed correlations that new measures of psychopathy display with the same external variables. Evidence that patterns produced by the PCL-R will generalize to a new psychopathy measure would follow from the demonstration that the new measure's pattern of associations with measures of criterion constructs is highly congruent with the pattern obtained by the PCL-R. Thus, in the present study we used the PCL-R's correlations with measures of criterion constructs as a basis for computing contrast values (λs) with which to compare *Zr* transformed correlations obtained using the LPSP and PPI.

The second adaptation of these measures was computational. The process Westen and Rosenthal (2003) suggested for computing $r_{\text{contrast-CV}}$ involves calculating (a) *Z* scores, (b) exact *p* values for those *Z* scores, (c) *t* scores for those exact *p* values, and (d) *r*, based on those *t* scores. In our study, four of eight *Z* scores and *df* (degrees of freedom) values fell out of range of this process (at step "b" or "c" above, applying SAS, STATA, Systat, Excel, and programmable scientific calculators). Thus, to convert *Z* scores into *t* values (to move from step "a" to step "c" above), we used a formula suggested by Robert Rosenthal (personal communication, April 20, 2008 in Rosenthal & Rosnow, 1991, p. 591):

$$Z = [df \log_e(1 + t^2/df)]^{1/2} [1 - 1/2df]^{1/2}.$$

The formula was written into an Excel program (available from the third author) that allows the user to solve for *t* and yields r_{contrast}. For the values that were within range of the process originally recommended by Westen and Rosenthal (2003), this formula and program yielded the same r_{contrast} values. In other words, the computational adaptation of the measures yields the same results.

Results

Zero-Order Correlations Among Psychopathy Measures

Total scores from each self-report psychopathy measure correlated significantly with the PCL-R Total score: $r = .30$ ($p < .001$) for the LPSP, and $r = .43$ ($p < .001$) for the PPI. A test for the difference between dependent correlations (Cohen & Cohen, 1983) revealed that these correlations differ significantly, t (1475 *df*) = –7.01, $p < .001$.

Correlations among the factor scores for all three psychopathy measures are shown in Table 1. Convergent validity for the self-report measures' factor scores was assessed by examining their associations with the corresponding PCL-R factor scores. Convergent validity was demonstrated for PPI-I and the LPSP Primary scale, which were moderately but significantly correlated with PCL-R Factor 1. The PPI-I and LPSP Primary scale correlations with PCL-R Factor 1 did not differ significantly, t (1469) = .60, *ns* (not significant). Convergent validity was also demonstrated for PPI-II and the LPSP Secondary scale, both of which were moderately and significantly correlated with PCL-R Factor 2,

Table 1. Correlations Between Factor Scores From Three Psychopathy Measures

		PCL-R	PPI		LPSP	
		Factor 2	PPI-I	PPI-II	Primary	Secondary
PCL-R	Factor1	.49	.25	.17	.23	.06
	Factor 2		.16	.39	.29	.29
PPI	PPI-I			–.06	.13	–.25
	PPI-II				.62	.70

Note: PCL-R = Hare Psychopathy Checklist–Revised; PPI = Psychopathic Personality Inventory; PPI-I = Fearless Dominance factor; PPI-II = Impulsive Antisociality factor; LPSP = Levenson Primary and Secondary Psychopathy scales. *N*s range from 1,472 to 1,484. All correlations are significant at $p = .02$ or greater (two-tailed).

with a significantly stronger association for the PPI-II, $t(1469) = 5.36$, $p < .001$.

Discriminant validity was assessed by examining the self-report measures' factor scores correlations with the noncorresponding PCL-R factor scores. Discriminant validity was observed for PPI-I, whose correlation with PCL-R Factor 1 was significantly stronger than its correlation with PCL-R Factor 2, $t(1469) = 2.45$, $p < .05$. However, the LPSP Primary scale demonstrated poor discriminant validity, as its association with PCL-R Factor 2 was significantly *higher* than its association with PCL-R Factor 1, $t(1471) = -2.41$, $p < .05$. Discriminant validity was demonstrated for PPI-II, whose correlation with PCL-R Factor 2 was significantly higher than its correlation with PCL-R Factor 1, $t(1469) = -6.26$, $p < .001$, and for the LPSP Secondary scale, whose association with PCL-R Factor 2 was significantly higher than its association with PCL-R Factor 1, $t(1470) = -9.25$, $p < .001$.

In summary, convergent validity was demonstrated for the LPSP and PPI total and factor scores via significant positive correlations with the corresponding PCL-R indices; the Total and Factor 2 score correlations (but not the Factor 1 correlations) were significantly higher for the PPI than for the LPSP. Discriminant validity was observed for PPI Factors 1 and 2 and for LPSP Factor 2; however, the LPSP Primary scale correlated more strongly with Factor 2 than with Factor 1 of the PCL-R, thus demonstrating poor discriminant validity.

Incremental Utility of the PPI and LPSP in Predicting PCL-R Scores

We used hierarchical multiple regression analyses to examine the incremental utility of the PPI and LPSP above and beyond one another for predicting PCL-R total scores. Entering LPSP total scores following PPI total scores yielded a nonsignificant increase in variance for predicting PCL-R total scores, $F_{change}(1, 1468) = .00$, ns, $R^2_{change} = .00$. In contrast, entering PPI total scores following LPSP total scores yielded a significant increase in variance for predicting PCL-R total scores, $F_{change}(1, 1468) = 174.18$, $p < .001$, $R^2_{change} = .10$.

In addition, we examined the incremental utility of PPI-I and the LPSP Primary scale above and beyond one another for predicting PCL-R Factor 1 scores, and the incremental utility of PPI-II and the LPSP Secondary scale above and beyond one another for predicting PCL-R Factor 2 scores. Entering LPSP Primary scores following PPI-I scores yielded a significant increase in variance for predicting PCL-R Factor 1 scores, $F_{change}(1, 1469) = 67.21$, $p < .001$, $R^2_{change} = .04$. Similarly, entering PPI-I scores following LPSP Primary scores yielded a significant and comparable increase in variance for predicting PCL-R Factor 1 scores, $F_{change}(1, 1469) = 81.09$, $p < .001$, $R^2_{change} = .05$. In the second set of analyses, entering LPSP Secondary scores following PPI-II scores did not yield a significant increase in variance for predicting PCL-R Factor 2 scores, $F_{change}(1, 1462) = .57$, ns, $R^2_{change} = .00$. In contrast, entering PPI-II scores following LPSP Secondary scores did yield a significant increase in variance for predicting PCL-R Factor 2 scores, $F_{change}(1, 1462) = 120.75$, $p < .001$, $R^2_{change} = .07$.

Similarity of PPI, LPSP, and PCL-R Correlations With External Variables

Correlations for the Total and factor scores of the PPI, LPSP, and PCL-R with the 35 criterion measures[4] described above (Measures; see also the legend for Figure 1) were used to compute two summary indices, $r_{alerting\text{-}CV}$ and $r_{contrast\text{-}CV}$,[5] to estimate the construct validity of the PPI and LPSP. For comparisons of associations of these measures' total scores, "Factor 1" scores, and "Factor 2" scores with external measures, the corresponding PCL-R score correlations were used to compute the λ values that constitute the appropriate contrast weights. As a "control," we also computed $r_{alerting\text{-}CV}$ and $r_{contrast\text{-}CV}$ for the Schizophrenia (SCZ) scale of the PAI. Unlike the PPI and LPSP, the SCZ scale is theoretically largely independent of psychopathy; indeed, Cleckley (1964) described the psychopath as "free from signs or symptoms traditionally regarded as evidence of a psychosis" (p. 366). Thus, substantial positive values for $r_{alerting\text{-}CV}$ and $r_{contrast\text{-}CV}$ were expected for the PPI and LPS, but not for SCZ.

The results of these analyses are shown on the diagonal of Table 2. Results displayed in the upper panel reveal substantial similarity in the patterns of associations for the PCL-R and PPI total scores, $r_{alerting\text{-}CV} = .80$ and $r_{contrast\text{-}CV} = .88$. In contrast, weaker positive indices of association were obtained for the LPSP total score, $r_{alerting\text{-}CV} = .61$ and $r_{contrast\text{-}CV} = .76$. The corresponding indices obtained for the SCZ scale

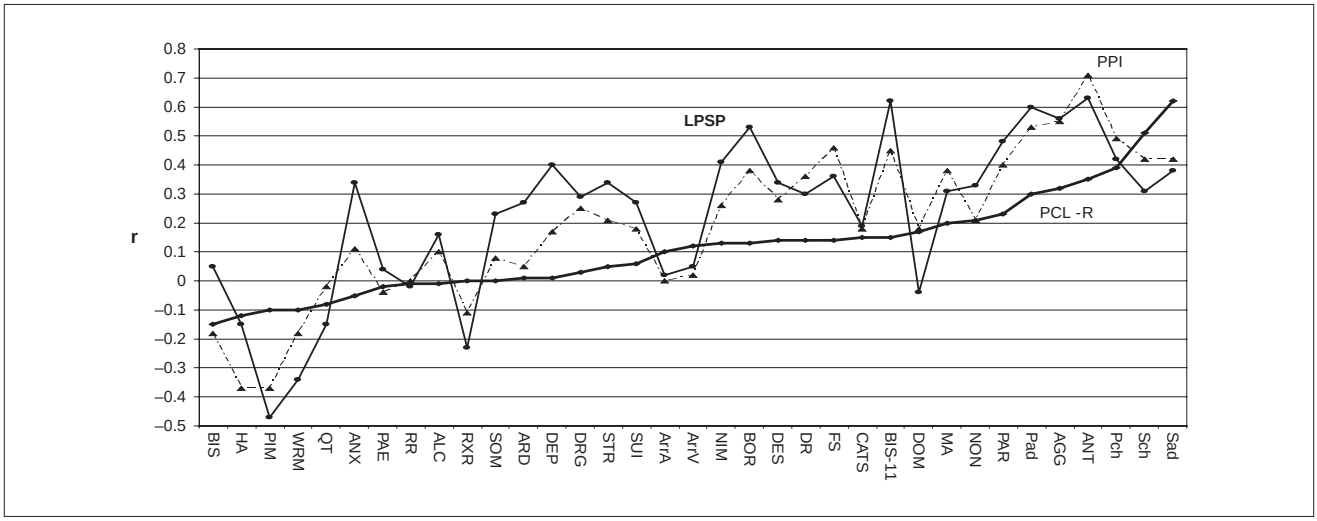

Figure 1. Correlations of the total scores from the PCL-R, PPI, and LPSP with 35 external variables
Note: PCL-R = Hare Psychopathy Checklist–Revised; PPI = Psychopathic Personality Inventory; BIS = Behavior Inhibition; HA = Harmavoidance; PIM = PAI (Personality Assessment Inventory) Positive Impression; WRM = PAI Warmth; QT = Quick Test IQ; ANX = PAI Anxiety; PAE = passive avoidance errors; RR = BAS Reward Responsivity scale; ALC = PAI Alcohol Problems; RXR = PAI Treatment Rejection; SOM = PAI Somatic Complaints; ARD = PAI Anxiety-Related Disorders; DEP = PAI Depression; DRG = PAI Drug Problems; STR = PAI Stress; SUI = PAI Suicidal Ideation; ArrA = any arrest; ArrV = arrest for violent offense; NIM = PAI Negative Impression; BOR = PAI Borderline Features; DES = Dissociative Experiences Scale; DR = BAS Drive; FS = BAS Fun Seeking; CATS = Child Abuse & Trauma; BIS-11 = Barratt Impulsivity Scale; DOM = PAI Dominance; MA = PAI Mania; NON = PAI Nonsupport; PAR = PAI Paranoia; Pad = adult symptoms on Personality Disorders Questionnaire 4+ (PDQ-4+) antisocial personality module; AGG = PAI Aggression; ANT = PAI Antisocial Features; Pch = child symptoms on PDQ-4+ antisocial personality module; Sch = child symptoms on Structured Clinical Interview for *DSM-IV* Axis II Personality Disorders (SCID-II) antisocial personality module; Sad = adult symptoms on SCID-II antisocial personality module.

Table 2. $r_{alerting\text{-}CV}$ and $r_{contrast\text{-}CV}$ for PPI and LPSP, Using λ based on PCL-R Correlations With 35 Measures to Define Criterion Validity Pattern

	PPI	LPSP	LPSP		LPSP	
PCL-R Criterion Index	Total	Total	PPI-I	Primary	PPI-II	Secondary
λ for PCL-R Total						
$r_{alerting\text{-}CV}$	.80	.61	—	—	—	—
$r_{contrast\text{-}CV}$	.88	.76	—	—	—	—
λ for PCL-R Factor 1						
$r_{alerting\text{-}CV}$	—	—	.58	.50	.18	.01
$r_{contrast\text{-}CV}$	—	—	.71	.59	.38	.07
λ for PCL-R Factor 2						
$r_{alerting\text{-}CV}$	—	—	.09	.85	.72	.58
$r_{contrast\text{-}CV}$	—	—	.14	.82	.93	.76

Note: PCL-R = Hare Psychopathy Checklist–Revised; PPI = Psychopathic Personality Inventory; LPSP = Levenson Primary and Secondary Psychopathy scales.

were $r_{alerting\text{-}CV} = -.18$ and $r_{contrast\text{-}CV} = .25$, indicating comparatively little similarity in patterns of associations between the SCZ and PCL-R measures. Thus, at the level of total score analyses, the pattern of external correlates for the PPI, and to a lesser degree that of the LPSP, is substantially similar to the pattern observed for the PCL-R.

Although somewhat weaker than the Total score results, indices displayed on the diagonal of Table 2 also reveal significant correspondence with the patterns of PCL-R criterion correlations for the Factor 1 (PPI-I and Primary, middle panel) and Factor 2 (PPI-II and Secondary, lower panel) indices from the self-report psychopathy measures. In both instances, the correspondence with patterns of criterion correlations surrounding the PCL-R factor scores is somewhat higher for the PPI factors than for the LPSP factors. Notably, for the PPI, indices of correspondence between its second factor (PPI-II) and Factor 2 of the PCL-R were higher than those between its first factor

(PPI-I) and Factor 1 of the PCL-R. This asymmetry was evident (to a lesser degree) for the LPSP factors.

Further information regarding the construct validity of the PPI and LPSP factors is provided in the off-diagonal cells of Table 2. The off-diagonal $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ values in the middle panel indicate correspondence between the patterns of criterion correlations for PPI-II and LPSP Secondary, respectively, with the pattern surrounding the *noncorresponding* PCL-R factor, Factor 1. These indices provide a perspective on the discriminant validity of the patterns of external variable correlations for PPI-II and LPSP Secondary. For both these factors, the evidence for discriminant validity is positive. For each measure, there is less congruence with the pattern of external correlations surrounding the *noncorresponding* PCL-R factor (Factor 1) than was obtained for the corresponding factor (PCL-R Factor 2, lower panel, on-diagonal).

The off-diagonal $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ values in the lower panel indicate correspondence between the patterns of criterion correlations for PPI-I and LPSP Primary, respectively, with the pattern surrounding the *noncorresponding* PCL-R factor, Factor 2. Here, satisfactory discriminant validity in terms of patterns of correlations with external variables is evident only for PPI-I. For PPI-I, the correspondence of its pattern of associations with external variables is much stronger with that surrounding PCL-R Factor 1 than PCL-R Factor 2. However, the opposite is true for the LPSP Primary scale. For the LPSP Primary scale $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ are higher for the *noncorresponding* PCL-R factor (Factor 2), suggesting that in its associations with external variables, the LPSP Primary scale functions more like a measure of PCL-R Factor 2 than Factor 1.

Smith (2005) cautioned against the blind interpretation of $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ values and provided several examples (table 1, p. 403) in which clearly aberrant (i.e., nonlinearly related) patterns of obtained correlations could inflate these estimates of a measure's construct validity. Although this is more likely to occur when the number of criterion measures is small (≤10), his concern illustrates the importance of visually inspecting for concordance between patterns of obtained and predicted correlations. Thus, Figure 1 displays the patterns of criterion correlations observed for the PPI and LPSP total scores plotted against the (ordered) correlations obtained for the PCL-R Total score.

Discussion

In this study, we used scores on the PCL-R as provisional benchmarks for the validation of the PPI and LPSP scales according to the standards articulated by Burisch (1984): namely, that

> Variable Y can be regarded as a criterion for a test X if and only if (a) they are measures of the same construct—that is, if they are theoretically or semantically fairly close—and (b) Y has a higher "status" than X—that is, Y is more "trustworthy." (p. 217)

Given that the PCL-R is currently the most extensively validated measure of psychopathy for offenders, reasonable grounds exist for asserting that it is presently a more "trustworthy" measure of psychopathy for this population than either the PPI or LPSP. If investigators or practitioners wish to consider using the PPI or LPSP in lieu of the PCL-R, comparisons of those measures with PCL-R performance provides important and relevant data for informing their judgments about which of them to use.

In this regard, our findings paint a reasonably consistent picture. Two key sets of results are worth highlighting. First, the PPI total score and PPI-II displayed statistically significant, albeit small to moderate, levels of incremental utility above and beyond the LPSP total and secondary scales for their target variables, namely PCL-R total and Factor 2 scores, respectively. In sharp contrast, the LPSP total and secondary scales displayed no incremental utility above and beyond the PPI total score and PPI-II for these variables. Both PPI-I and the LPSP primary scale exhibited statistically significant, albeit small, levels of incremental utility above and beyond each other for PCL-R Factor 1 scores. This suggests that the PPI, compared with the LPSP, provides better coverage of the PCL-R as a whole and its deviant lifestyle component, but not necessarily its interpersonal/affective component.

Second, the results of an adapted Westen and Rosenthal (2003) analysis suggest that the PPI's total score and its two factors better replicate the PCL-R's pattern of correlations with external variables than do the LPSP total score and its Primary and Secondary scales. Of particular concern with the LPSP was the finding that the pattern of external correlations for its Primary scale was closer to PCL-R Factor 2 than Factor 1.

These findings dovetail with others suggesting that the LPSP Primary scale does not adequately assess the core interpersonal and affective features of psychopathy (Lilienfeld & Fowler, 2006). Miller, Gaughan, and Pryor (2008) argued that

> The LSRP [referred to as the LPSP in the present investigation] F1 appears to result in a general personality profile that is as similar to the profile generated by the PCL-R F1 as those generated by the factor 1 scores of the PPI.

The present findings, however, challenge this contention. Instead, they demonstrate that, as gauged by its pattern of

external correlates, the concordance between the LPSP primary scale and PCL-R Factor 1 is somewhat weaker than that between PPI-I and PCL-R Factor 1. More important, they indicate that, unlike PPI-I, the LPSP primary scale is actually more similar to PCL-R Factor 2 than PCL-R Factor 1 in its pattern of external correlates. This latter finding suggests that the anomalous pattern of correlates for the LPSP primary scale cannot merely be attributed to the operationalization of the first psychopathy factor in the PPI and other self-report measures (cf. Miller et al., 2008). In fact, using both Westen and Rosenthal (2003) indices, the LPSP primary scale was more similar to PCL-R Factor 2 than was the LPSP Secondary scale, even though the latter scale was designed to map onto PCL-R Factor 2 (Levenson et al., 1995).

Nevertheless, it is important to emphasize that our findings do not imply that the PPI factors are isomorphic or even highly similar to their corresponding PCL-R factors, only that they bear notable commonalities in their pattern of relations with external criterion measures. Indeed, direct relations between these two sets of factors were well below levels expected for alternative measures of the same construct (i.e., below the product of their respective reliabilities), with the association between Factor 1 scores for the two instruments lower ($r = .25$) than that between scores for their respective Factor 2's ($r = .39$). Echoing prior findings (e.g., Malterer et al., in press; Patrick et al., 2006), these results indicate that the PPI factors, in particular PPI-I, are operationalizing psychopathy in a distinctly different fashion than the PCL-R factors.

This is due in part to the fact that the two inventories employ different assessment methods (i.e., self-report versus interviewer ratings); measures of the same construct assessed in different measurement domains are expected to correlate with one another only moderately (.4 to .6), rather than highly (.7 to .9; Campbell & Fiske, 1959). However, recent theoretical analyses (Patrick, Fowles, & Krueger, 2009) suggest substantive differences in aspects of psychopathy captured by PCL-R Factor 1 and PPI-I. The PPI-I subscales index social potency, fearlessness, and stress immunity, the latter two of which are features that are not explicitly assessed by the PCL-R. PPI-I appears to capture features that represent a "boldness" phenotype characterized by social dominance, emotional resiliency, and venturesomeness. In contrast, PCL-R Factor 1 assesses features such as lack of remorse or guilt and callousness/lack of empathy, thought to relate to a "meanness" phenotype that represents aggressive resource-seeking without regard for others. That the "boldness" and "meanness" constructs are only partially overlapping may explain the modest correlation between PPI-I and PCL-R Factor 1.

Future research will be needed to ascertain whether the PPI and its factors are more or less valid than the PCL-R and its factors for predicting theoretically relevant (e.g., poor passive avoidance learning; deficient startle modulation) and pragmatically important (e.g., treatment response) external criteria. With respect to pragmatic criteria, the PPI may be at a disadvantage for predicting violent and other criminal behavior because the PCL-R includes explicitly criminological items (e.g., juvenile delinquency and diversity of prior offending) that predict future offending behavior (Walters, 2003). Still, the PPI may prove to be useful in correctional settings, given that it has outperformed the PCL-R in predicting suicide-related behavior (Douglas et al., 2008) and disciplinary infractions in prison (Edens et al., 2008).

As one thoughtful reviewer noted, an arguably more appropriate "benchmark" for evaluating new measures of psychopathy would in principle be a pattern of associations with criterion measures established on the basis of theory. In this light, our selection of the PCL-R as a provisional benchmark for evaluating the construct validity of the PPI and LSPS means that our comparisons are necessarily pragmatic rather than theoretically grounded. That is, our analyses bear on the practically useful question of which self-report measure better approximates the PCL-R in its pattern of external correlates, but do not address the question of which measure better conforms to a pattern of correlates derived from theory. Nevertheless, because there are several competing etiological models of psychopathy (see Patrick, 2006), including a fearlessness model (Lykken, 1995), a response modulation model (Newman, Schmitt, & Voss, 1997), a violence inhibition model (Blair, Jones, Clark, & Smith, 1997), and a dual process model (Fowles & Dindo, 2006), generating consensual theoretical predictions regarding the correlates of psychopathy measures is far from straightforward. Although further light may be shed on psychopathy by applying this approach from specific theoretical vantage points, we regard our findings as a helpful first step toward elucidating the correlates of two widely used self-report psychopathy measures and ascertaining which measure might be more likely to produce associations with criterion measures similar to those of the PCL-R. That said, our choice of criterion measures was not arbitrary. As noted earlier (see Measures section), many of the criterion measures used here have been linked explicitly, either clinically or in theory, to psychopathy, and all have been used in prior studies of the external validity of psychopathy measures.

Our findings must be interpreted in light of at least five limitations. First is the use of the PCL-R scores and their associations with criterion measures as provisional benchmarks for the psychopathy construct. We believe that this choice is justified by the status of the PCL-R as currently the best-validated measure of psychopathy with offenders. Nevertheless, this assertion is epistemic, not ontological, in nature. That is, we do not wish to imply that the PCL-R will ultimately be shown to be a more valid measure of psychopathy than the PPI or LPSP scales, only that it is more

extensively validated with offenders at present. Moreover, the PCL-R's conceptualization and operationalization of psychopathy have recently been criticized on several grounds, including its inadequate coverage of potentially adaptive features of psychopathy, such as social efficacy and emotional resiliency (Lynam & Widiger, 2007; Patrick, 2006; Patrick et al., 2009) and its heavy emphasis on indicators involving criminal behavior (Andrade, 2008; Blackburn, 2007). Therefore, our findings using Westen and Rosenthal's (2003) framework to compare the PPI and LPSP could differ when using a psychopathy "benchmark" that operationalizes psychopathy in a different manner than the PCL-R.

Second, by definition, our analyses using Westen and Rosenthal's (2003) metrics were constrained by our selection of validating variables. Although we cast a broad net of variables spanning measures of personality, Axis I pathology, Axis II pathology, childhood abuse, passive-avoidance learning, intelligence, and criminal recidivism, our findings for the PPI and LPSP scales might have differed had we administered alternative validity indicators. Future research will be needed to determine whether our findings hold up when alternative validating variables, particularly those that may serve as "endophenotypic" markers (Gottesman & Gould, 2003; Waldman, 2005) of basic processes associated with psychopathy, such as brain imaging or psychophysiological response measures, are included.

Third, the pattern of correlations observed in this study (and hence, $r_{alerting}$ and $r_{contrast}$) reflect not only construct variance, but also method variance. Like most of the criterion variables, the PPI and LPSP are self-report measures, whereas the PCL-R reflects an observer's ratings. It is probable that method variance inflated many of the correlations for the PPI and LPSP, relative to the PCL-R. Nevertheless, this concern applies with equal force to the PPI and LPSP and therefore cannot explain the differences in the extent to which their patterns of criterion correlations mirrored those of the PCL-R.

Fourth, both the PPI and LPSP scales were developed primarily for use in nonclinical and noncriminal samples (Lilienfeld & Fowler, 2006), although they have since been extended for use in prisons and psychiatric settings. As a consequence, it is unknown whether the present findings would extend to the more high-functioning samples for which these measures were initially designed.

Finally, the current study relied on the most frequently used factor structures, namely, a two-factor structure, for each of the three psychopathy instruments. This approach is justifiable given that it bears clear implications for the large body of literature based on these measures; moreover, it facilitates comparisons of their putatively corresponding factors. Nevertheless, there is controversy about the factor structure of each measure (see Note 1) and the findings reported here might differ if alternative structures were used to compare these measures.

These important caveats notwithstanding, our findings provide more compelling support for the construct validity of the PPI as a measure of psychopathy than the LPSP, at least in offender samples. They also suggest that research interpretations of the LPSP Primary scale in terms of the core interpersonal and affective features of psychopathy need to be tempered in view of this scale's questionable construct validity. Finally, the present results remind us of the hazards of the "jangle fallacy," the error of assuming that measures that carry the same name are necessarily assessing the same construct (Block, 1995; Thorndike, 1903). Not only are the PPI and LPSP factors not isomorphic with their corresponding PCL-R factors, they differ from each other in their pattern of external correlates. Consumers of the psychopathy literature must bear these differences in mind when interpreting divergent findings derived from alternative measures of psychopathy.

Acknowledgments

We would like to acknowledge the extensive assistance provided by Robert Rosenthal in the statistical analyses involving computations of $r_{contrast\text{-}CV}$. We acknowledge and appreciate the assistance and cooperation of the following agencies in collecting data for this research; however, none of the opinions or conclusions expressed in this article reflects any official policy or position of any of these institutions: Drug Abuse Comprehensive Coordinating Office (DACCO), Tampa, FL; Florida Department of Corrections; Gateway Foundation, Huntsville, TX; Nevada Department of Prisons; Odyssey House, Salt Lake City, UT; Operation PAR, Pinellas Park, FL; Oregon Department of Corrections; Texas Department of Criminal Justice–Institutional Division; Utah Department of Corrections; Volunteers of America, Portland, OR; Westcare, NV.

Declaration of Conflicting Interests

The authors declared a potential conflict of interest as follows: Scott Lilienfeld is the co-author of the Psychopathic Personality Inventory–Revised.

Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article:

National Institute of Mental Health (Grant #RO1 MH63783 -01A1)

Notes

1. There is controversy regarding the two-factor models of the PCL-R (see, e.g., Cooke & Michie, 2001), the PPI (see, e.g., Neumann et al., 2008), and the LPSP (see, e.g., Brinkley et al., 2008). However, given the objectives of the present study, we opted to score each measure according to its most widely used and researched structure.
2. The number of cases used in correlational analyses varied due to missing data or, in the case of our recidivism measure (see

Measures section), because data were collected only on that subsample of individuals released to the community; *n*s ranged from 1,016 (PCL-R Factor 2 correlation with recidivism outcomes) to 1,603 (associations among PAI scales).

3. One thoughtful reviewer expressed concerns that we included in our analyses those participants to whom RAs (research assistants) read aloud items from the self-report measures. To evaluate the potential impact of including these cases, we recalculated correlation coefficients between our 35 criterion measures (see Figure 1) and the Total, Factor 1, and Factor 2 scores for the three psychopathy measures after removing these 44 participants. Of these 315 correlations, 61% were identical and 37% differed by ±.01. Of the 5 remaining correlations, 4 differed by only ±.02 and 1 differed by .05. Given the negligible impact on our results, we retained these 44 cases in the analyses.
4. A table of these correlations may be requested from the first author. One thoughtful reviewer noted that some observed correlations for the PCL-R with criterion measures were smaller in magnitude than he or she might have expected from the previous literature. For example, small correlations were obtained for the PCL-R Total and factors scores with number of errors on the go/nogo passive avoidance task (*r*s range from .02 to .03), with general recidivism (*r*s range from .06 to .12), and with violent recidivism (*r*s range from .06 to .11). Although we agree that these correlations are lower *on average* than what has been reported, they are within the range correlations that have been reported previously. Nevertheless, to assess whether inclusion of these small correlations in our analyses may have produced misleading results, we recalculated the $r_{\text{alerting-CV}}$ indices for the PPI and LPSP total and factor scores with their corresponding PCL-R indices, after removing these three variables. In no instance did the value of $r_{\text{alerting-CV}}$ change by more than .02. We interpret these results as indicating that the inclusion of these variables in our analyses did not produce misleading results regarding similarities in the patterns of correlations with external variables.
5. The formulae provided by Westen and Rosenthal (2003) for the computations required to obtain $r_{\text{contrast-CV}}$ assume equal *N* of subjects for all criterion correlations. As a result of missing data for various participants, *N*s were unequal for our criterion correlations. Thus, as recommended by Robert Rosenthal (personal communication, December 17, 2007), our computations used the harmonic mean, n_h for $k = 35$ criterion variables, using the formula

$$n_h = \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_k}}.$$

This yielded the following n_h values for our computations: PPI Total, n_h = 1,466; PPI-I, n_h = 1,466; PPI-II, n_h = 1,464; LPSP Total, n_h = 1,489; LPSP Primary, n_h = 1,490; LPSP Secondary, n_h = 1,489; SCZ, n_h = 1,519.

References

Ammons, R. B., & Ammons, C. H. (1962). The Quick Test (QT): Provisional manual. *Psychological Reports, 11*, 111-161.

Andrade, J. T. (2008). The inclusion of antisocial behavior in the construct of psychopathy: A review of the research. *Aggression and Violent Behavior, 13*, 328-335.

Barratt, E. S. (1994). Impulsiveness and aggression. In J. Monahan & H. J. Steadman (Eds.), *Violence and mental disorder*: *Developments in risk assessment* (pp. 61-80). Chicago: University of Chicago Press.

Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the Psychopathic Personality Inventory: Validity and implications for clinical assessment. *Psychological Assessment, 15*, 340-350.

Benning, S. D., Patrick, C. J., Salekin, R. T., & Leistico, A. R. (2005). Convergent and discriminant validity of psychopathy factors assessed via self-report: A comparison of three instruments. *Assessment, 12*, 270-289.

Berardino, S. D., Meloy, J. R., Sherman, M., & Jacobs, D. (2005). Validation of the Psychopathic Personality Inventory on a female offender sample. *Behavioral Sciences & the Law, 23*, 819-836.

Blackburn, R. (2007). Personality disorder and antisocial deviance: Comments on the debate on the structure of the Psychopathy Checklist–Revised. *Journal of Personality Disorders, 21*, 142-159.

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology, 34*, 192-198.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187-215.

Borsboom, D., Mellenbergh, G. J., & Heerden, J. V. (2004).The concept of validity. *Psychological Review, 111*, 1061-1071.

Brinkley, C. A., Diamond, P. M., Magaletta, P. R., & Heigel, C. P. (2008). Cross-validation of Levenson's psychopathy scale in a sample of federal female inmates. *Assessment. 15*, 464-482.

Brinkley, C. A., Schmitt, W. A., Smith, S. S., & Newman, J. P. (2001). Construct validity of a self-report psychopathy scale: Does Levenson's self-report psychopathy scale measure the same construct as Hare's Psychopathy Checklist–Revised? *Personality and Individual Differences, 31*, 1021-1038.

Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist, 39*, 214-227.

Cale, E. M., & Lilienfeld, S. O. (2002). Sex differences in psychopathy and antisocial personality disorder: A review and integration. *Clinical Psychology Review, 22*, 1179-1207.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait, multi-method matrix. *Psychological Bulletin, 56*, 81-105.

Carlson, E. B., & Putnam, F. W. (1993). An update on the Dissociative Experiences Scale. *Dissociation, 6*, 16-27.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319-333.

Cleckley, H. (1964). *The mask of sanity*. St. Louis, MO: Mosby. (Original work published 1941)

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

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

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Douglas, K. S., Guy, L. S., Edens, J. F., Boer, D. P., & Hamilton, J. (2007). The Personality Assessment Inventory as a proxy for the Psychopathy Checklist–Revised: Testing the incremental validity and cross-sample robustness of the Antisocial Features Scale. *Assessment, 14*, 255-269.

Douglas, K., Lilienfeld, S. O., Skeem, J., Poythress, N. G., Edens, J. F., & Patrick, C. J. (2008). Relation of psychopathic and antisocial traits to suicide-related behavior among offenders. *Law and Human Behavior, 32*, 511-525.

Edens, J. F., Hart, S. D., Johnson, S. W., Johnson, J. K., & Olver, M. E. (2000). Use of the Personality Assessment Inventory to assess psychopathy in offender populations. *Psychological Assessment, 12*, 132-139.

Edens, J. F., Poythress, N. G., Lilienfeld, S. O., & Patrick, C. J. (2008). A prospective comparison of two measures of psychopathy in the prediction of institutional misconduct. *Behavioral Sciences & the Law, 26*, 529-541.

Edens, J. F., & Ruiz, M. (2005). *Personality Assessment Inventory Interpretive Report for Correctional Settings (PAI-CS) professional manual*. Odessa, FL: Psychological Assessment Resources.

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. (1996). *User's guide for the structured clinical interview for DSM-IV Axis II Personality Disorders* (SCID-II, Version 2.0, July 1996 Final Version). New York: Biometrics Research Department, New York State Psychiatric Institute.

Fowles, D. C. (1980). The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology, 17*, 87-104.

Fowles, D. C., & Dindo, L. (2006). A dual-deficit model of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 14-34). New York: Guilford Press.

Gottesman, I. I., & Gould, T. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636-645.

Gray, J. A. (1987). Perspectives on anxiety and impulsivity: A commentary. *Journal of Research in Personality, 21*, 493-509.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences, 1*, 111-119.

Hare, R. D. (1991). *The Hare Psychopathy Checklist–Revised*. Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D. (2003). *The Hare Psychopathy Checklist–Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Harpur, T., Hare, R., & Hakstian, R. (1989). A two-factor conceptualization of psychopathy: Construct validity and implications for assessment. *Psychological Assessment, 1*, 6-17.

Hyler, S. E. (1994). *PDQ-4+ Personality Diagnostic Questionnaire*. New York: New York State Psychiatric Institute.

Johns, J. (1997). *The basic reading inventory* (7th ed.). Dubuque, IA: Kendall/Hunt.

Karpman, B. (1948). The myth of the psychopathic personality. *American Journal of Psychiatry, 104*, 523-534.

Karpman, B. (1949). Psychopathy as a form of social parasitism: A comparative biological study. *Journal of Clinical Psychopathology, 10*, 160-194.

Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations: A survey of experts. *Professional Psychology: Research and Practice, 34*, 491-498.

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*, 151-158.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14*, 17-38.

Lilienfeld, S. O., & Andrews, B. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66*, 488-524.

Lilienfeld, S. O., & Fowler, K. A. (2006). The self-report assessment of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 107-132). New York: Guilford Press.

Lykken, D. (1995). *The antisocial personalities*. Hillsdale, NJ: Lawrence Erlbaum.

Lynam, D. R., & Widiger, T. A. (2007). Using a general model of personality to identify the basic elements of psychopathy. *Journal of Personality Disorders, 21*, 160-178.

Malterer, M., Lilienfeld, S. O., Neumann, C. S., & Newman, J. P. (in press). Construct validity of the Psychopathic Personality Inventory with offenders and community samples. *Assessment*.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35*, 1012-1027.

Miller, J. D., Gaughan, E. T., & Pryor, L. R. (2008). The Levenson Self-report Psychopathy Scale: An examination of the personality traits and disorders associated with the LSRP factors. *Assessment, 15*, 450-463.

Morey, L. C. (1991). *Personality Assessment Inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Neumann, C. S., Malterer, M. B., & Newman, J. P. (2008). Factor structure of the Psychopathic Personality Inventory (PPI): Findings from a large incarcerated sample. *Psychological Assessment, 20*, 169-174.

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology, 95*, 257-263.

Newman, J. P., & Schmitt, W. A. (1998). Passive avoidance in psychopathic offenders: A replication and extension. *Journal of Abnormal Psychology, 107*, 527-532.

Newman, J. P., Schmitt, W., & Voss, W. (1997). The impact of motivationally neutral cues on psychopathic individuals: Assessing the generality of the response modulation hypothesis. *Journal of Abnormal Psychology, 106*, 563-575.

Patrick, C. J. (1994). Emotion and psychopathy: Startling new insights. *Psychophysiology, 31*, 319-330.

Patrick, C. J. (2001). Emotional processes in psychopathy. In A. Raine & J. Sanmartin (Eds.), *Violence & psychopathy* (pp. 57-77). New York: Kluwer/Plenum.

Patrick, C. J. (2006). Back to the future: Cleckley as a guide to the next generation of psychopathy research. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 605-618). New York: Guilford Press.

Patrick, C. J., Edens, J. F., Poythress, N. G., Lilienfeld, S. O., & Benning, S. (2006). Construct validity of the Psychopathic Personality Inventory two-factor model with offenders. *Psychological Assessment, 18*, 204-208.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology, 21*, 913-938.

Porter, S. (1996). Without conscience or without active conscience? The etiology of psychopathy revisited. *Aggression and Violent Behavior, 1*, 179-189.

Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (1998). Criterion-related validity of the Psychopathic Personality Inventory in a criminal population. *Psychological Assessment, 10*, 426-430.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Sanders, B., & Giolas, M. H. (1991). Dissociation and childhood trauma in psychologically disturbed adolescents. *American Journal of Psychiatry, 148*, 50-54.

Sandoval, A., Hancock, D., Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (2000). Construct validity of the Psychopathic Personality Inventory in a correctional sample. *Journal of Personality Assessment, 74*, 262-281.

Skeem, J., Johansson, P., Andershed, H., Kerr, M., & Louden, J. E. (2007). Two subtypes of psychopathic violent offenders that parallel primary and secondary variants. *Journal of Abnormal Psychology, 116*, 385-409.

Smith, G. (2005). On construct validity: Issues of method and measurement. *Psychological Assessment, 17*, 396-408.

Swogger, M. T., & Kosson, D. S. (2007). Identifying subtypes of criminal psychopaths: A replication and extension. *Criminal Justice and Behavior, 34*, 953-970.

Tellegen, A. (1982). *Brief manual for the Multidimensional Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.

Thorndike, E. L. (1903). *Educational psychology*. New York: Lemcke & Buechner.

Vassileva, J., Kosson, D. S., Abramowitz, C., & Conrod, P. (2005). Psychopathy versus psychopathies in classifying criminal offenders. *Legal and Criminological Psychology, 10*, 27-43.

Waldman, I. D. (2005). Statistical approaches to complex phenotypes: Evaluating neuropsychological endophenotypes for attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*, 1347-1356.

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*, 541-558.

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology, 84*, 608-618.