Since the introduction of the Psychopathy Checklist-Revised in 1991 (PCL-R), numerous studies have examined the reliability and validity of this assessment instrument in samples of institutionalized and noninstitutionalized females. We reviewed these data to determine the current state of PCL-R psychopathy assessment in females and to guide a discussion of the clinical utility of the PCL-R in female samples. Here we discuss issues generated by existing data and offer specific research suggestions. This review shows that there is support for the measure’s reliability in women and modest support for its validity. However, studies of the predictive utility and the factor structure of the PCL-R in women are required to demonstrate the generalizability of PCL-R psychopathy to women.

**Key words:** assessment, offenders, psychopathy, Psychopathy Checklist-Revised, women. [Clin Psychol Sci Pract 8:117–132, 2001]

Psychopathic individuals are characterized by a collection of antisocial behaviors, anomalous affective experiences, and callous, self-serving attitudes. Their behaviors include criminal activity ranging from drug dealing to assault. Their interpersonal interactions can be characterized by both superficial charm and a striking callousness toward others, often including a lack of remorse for the negative consequences of their behaviors and a lack of empathy. Despite the fact that historically the behaviors and attitudes of psychopathic individuals have taken a serious financial, psychological, and emotional toll on the individuals around them (see Millon, 1981), it was not until the first publication of Cleckley’s book, *The Mask of Sanity*, in 1941 that interest in studying these individuals blossomed. Cleckley’s (1976) book provided the first truly comprehensive descriptions of psychopathic individuals, from which Cleckley distilled 16 core traits of psychopathy (see Table 1). In the years following the formulation of these “Cleckley criteria,” numerous attempts were made to transform these clinical descriptions into a reliable measure for the assessment of psychopathy (see Hare & Cox, 1978, for a review).

The introductions of the Psychopathy Checklist (PCL; Hare, 1980) and the Psychopathy Checklist Revised (PCL-R; Hare, 1991) exemplify such attempts. Derived in large part from the original Cleckley criteria, the PCL-R is composed of 20 items that are individually scored as 0 (not present), 1 (may be present) or 2 (definitely present) (see Table 2). Ratings for each item are based on an interview with the participant as well as collateral information obtained from prison files. The checklist is composed of two moderately correlated factors. Factor 1 assesses callous disregard for the feelings and rights of others and includes such items as lack of guilt/remorse, callous/lack of empathy, and shallow affect. Factor 2 assesses persistent antisocial behavior and includes juvenile delinquency and criminal versatility (Harpur, Hakstian, & Hare, 1988).

The field of psychopathy was radically altered by Hare’s introduction of the PCL (1980) and the PCL-R (1991). Now considered the state-of-the-art measure of psychopathy (Fulero, 1995), the PCL-R has been widely adopted by researchers and clinicians in the field, allowing for the systematic classification and study of psychopathic individuals (Hare, 1996). Although the PCL-R includes numerous behavioral criteria specific to incarcerated offenders (e.g., “revocation of conditional release,”
than individuals with low scores (Serin, 1994), are prone to committing instrumental violence (Cornell, Warren, Hawk, Stafford, Oram, & Pine, 1996), and are more likely to recidivate violently (Hare, 1996; Hemphill, Templeman, Wong, & Hare, 1998; Serin, Peters, & Barbaree, 1990). It has also been suggested that high scorers are less successful in treatment programs (Ogloff, Wong, & Greenwood, 1990).

Research conducted using the PCL-R has increased our understanding of the causes and consequences of psychopathy. However, because most PCL-R research has been conducted with incarcerated, Caucasian males, it is difficult to know if the hypotheses and findings generated from studies of this group can be generalized to more diverse samples. The need to demonstrate empirically the generalizability of findings from one population to another is part of the scientific process. As Sue (1999) notes: “Many principles can be applied to different populations. Problems occur when the assumption of generality is made. Generality is a phenomenon that should be empirically tested” (p. 1074). Consistent with these scientific principles, the PCL-R manual (Hare, 1991) states explicitly that use of the instrument in clinical settings should be limited to “those populations in which it has been fully validated. For the present, this means adult male forensic populations” (p. 5).

In the years since Hare published the PCL-R and wrote this statement, a number of researchers have examined the reliability and validity of the PCL-R in samples of female offenders and nonoffenders. Recognizing that validation is an ongoing process, we would like to clarify the current state of the PCL-R in female samples and to suggest empirical questions that researchers in this area may wish to pursue. By reviewing existing empirical data and the various methodological and theoretical issues that these data have generated, we hope to provide a better understanding of the psychopathic female as she is assessed by the Psychopathy Checklist-Revised. Further, to place this review in context, we briefly review gender differences and related issues in the diagnosis of antisocial personality disorder and conduct disorder. These literatures provide potential models for interpreting existing data on gender differences in PCL-R–assessed psychopathy and for guiding researchers in the psychopathy area.

**GENDER AND ANTISOCIAL BEHAVIOR**

Interest in evaluating the extent of gender differences in PCL-R–assessed psychopathy is consistent with the

---

**Table 1.** The 16 Cleckley criteria for psychopathy

| 1. Superficial charm and good “intelligence” |
| 2. Absence of delusions or other signs of irrational thinking |
| 3. Absence of “nervousness” or psychoneurotic manifestations |
| 4. Unreliability |
| 5. Untruthfulness and insincerity |
| 6. Lack of remorse or shame |
| 7. Inadequately motivated antisocial behavior |
| 8. Poor judgment and failure to learn by experience |
| 9. Pathologic egocentricity and incapacity for love |
| 10. General poverty in major affective reactions |
| 11. Specific loss of insight |
| 12. Unresponsiveness in general interpersonal relations |
| 13. Fantastic and uninviting behavior with drink and sometimes without |
| 14. Suicide rarely carried out |
| 15. Sex life impersonal, trivial, and poorly integrated |
| 16. Failure to follow any life plan |

---

**Table 2.** The 20 PCL-R items

| 1. Glibness/superficial charm |
| 2. Grandiose sense of self-worth |
| 3. Need for stimulation/proneness to boredom |
| 4. Pathological lying |
| 5. Conning/manipulative |
| 6. Lack of remorse or guilt |
| 7. Shallow affect |
| 8. Callous/lack of empathy |
| 9. Parodic lifestyle |
| 10. Poor behavioral controls |
| 11. Promiscuous sexual behavior |
| 12. Early behavioral problems |
| 13. Lack of realistic, long-term goals |
| 14. Impulsivity |
| 15. Irrresponsibility |
| 16. Failure to accept responsibility for own actions |
| 17. Many short-term marital relationships |
| 18. Juvenile delinquency |
| 19. Revocation of conditional release |
| 20. Criminal versatility |

“criminal versatility”) that limit the measure’s utility in noninstitutionalized samples, no other measure has been used in research to the same degree that the PCL-R has been. Studies incorporating PCL-R assessments have helped researchers in both forensic and clinical settings to understand better both the causes of psychopathic individuals’ attitudes and behaviors and their negative impact on society. For example, research shows that incarcerated individuals scoring high on the PCL-R may not process the emotional content of words or pictures in the same way as individuals with low scores (e.g., Patrick, 1994; Patrick, Cuthbert, & Lang, 1994; Williamson, Harpur, & Hare, 1991) and that psychopathic individuals have difficulty reflecting on behaviors that have resulted in punishment (e.g., Newman, Patterson, & Kosson, 1987). In addition, research has shown that individuals with high PCL-R scores commit more violent criminal offenses than individuals with low scores (Serin, 1994), are prone to committing instrumental violence (Cornell, Warren, Hawk, Stafford, Oram, & Pine, 1996), and are more likely to recidivate violently (Hare, 1996; Hemphill, Templeman, Wong, & Hare, 1998; Serin, Peters, & Barbaree, 1990). It has also been suggested that high scorers are less successful in treatment programs (Ogloff, Wong, & Greenwood, 1990).
movement toward examining gender differences in psychopathology more generally (e.g., Gove & Tudor, 1973; Myers et al., 1984; Robins et al., 1984). Particular attention has been paid to syndromes that include antisocial or aggressive behaviors. Reviews and investigations of gender differences in the diagnosis, expression, and etiology of antisocial syndromes, such as conduct disorder (CD) in adolescents, and antisocial personality disorder (ASPD) in adults parallel the growing literature on gender differences in psychopathy and have addressed many of the issues that psychopathy researchers are only beginning to explore.

One of the most reliable and pivotal findings in these areas is that, in the general population, more males than females are diagnosed with both CD (Hartung & Widiger, 1998; Robins, 1999; Zahn-Waxler, 1993) and ASPD (American Psychiatric Association, 1987; Rutherford, Alterman, Cacciola, & Snider, 1995). This gender gap in the prevalence of these disorders has caused investigators to question whether the criteria for CD and ASPD are adequate for identifying females with these disorders. Some researchers have argued that the criteria may not be sufficient. For example, Zoccolillo, Tremblay, and Vitaro (1996) examined the rates of CD in a sample of preadolescent girls who had exhibited persistent and pervasive antisocial behavior beginning in kindergarten. The authors found that only 6% of the girls who were still exhibiting high levels of disruptive and antisocial behavior met DSM-III-R criteria for conduct disorder at age 10. The authors argued that the criteria for CD may be too restricted to males and may miss many preadolescent girls who have been exhibiting persistent antisocial behavior from an early age (Zoccolillo et al., 1996). In further support of this argument, the authors found that when the CD criteria were modified slightly, the percentage of persistently antisocial girls who met the CD criteria rose to 35%, while only 1% of the nonpersistently antisocial girls met the new criteria (Zoccolillo et al., 1996). In short, by altering the criteria slightly, the authors found that the diagnosis increased in sensitivity but did not lose specificity.

As a result of such studies, it was proposed that the DSM-III-R criteria for CD were not appropriate for girls. The solution, according to this argument, would be to employ gender-specific criteria for CD consistent with differences in the base rates of aggression and criminality in boys and girls (Zoccolillo, 1993). Although this suggestion has not been accepted universally (e.g., Zahn-Waxler, 1993), it does highlight the argument that externalizing disorders may be expressed differently across gender and that, as a result, current diagnostic criteria may fail to identify affected females.

As in CD, researchers in ASPD have expended considerable effort to determine the specific ways in which the diagnostic criteria for ASPD may be gender biased. These studies attempt to pinpoint specific gender differences in the presentation and correlates of ASPD. For example, Goldstein, Powers, McCusker, and Mundt (1996) examined gender differences in the expression of DSM-III-R ASPD in a sample of male and female drug abusers. The authors found gender differences in the behavioral manifestations of ASPD in both childhood and adulthood. In childhood, women with ASPD had run away more often than men, but had less often used weapons in fights, been cruel to animals, set fires, and perpetrated vandalism. In adulthood, women with ASPD were more often irresponsible as parents and in financial matters, engaged in prostitution, made money finding customers for prostitutes, were physically violent against sex partners and children, and failed to plan ahead.

ASPD has also been found to relate differentially to substance abuse in men and women. Studies of this relationship have shown that ASPD is related more strongly to both alcoholism (Hesselbrock & Hesselbrock, 1997) and cocaine dependence (Weiss, Martinez-Raga, Griffin, Greenfield, & Hufford, 1997) in men than in women.

A number of theorists have addressed the possibility of gender-specific manifestations of psychopathology by examining the within-family associations of various disorders. For example, there is considerable familial overlap among ASPD, somatization disorder, and histrionic personality disorder (Hamburger, Lilienfeld, & Hogben, 1996; Lilienfeld, Van Valkenburg, Larnzt, & Akiskal, 1986; Paris, 1997), a finding that researchers have proposed may represent the same underlying diathesis being given alternative expressions according to gender.

The distinction between gender differences in the manifestation of a process versus gender differences in the process itself is clearly relevant to studies of gender differences in psychopathy. Further, the parallel between the gender issues addressed in ASPD and CD and gender issues in psychopathy enables researchers to consider the gender differences we see in psychopathy within a related, existing literature. Consider, for example, the question of whether researchers should expect differences in the base rate of psychopathy across men and women. Research on CD and ASPD has shown that there are potential differ-
ences in the expression of these disorders across gender and that there are gender differences in the prevalence of these disorders. These findings might lead us to expect gender differences in the base rates of psychopathy.

Other studies, however, have not found differences in the prevalence of either ASPD or CD in certain samples of males and females (e.g., Kokkevi & Stefamos, 1995; Lewis, Helzer, Cloninger, Croughan, & Whitman, 1982; Rounds-Bryant, Kristiansen, Fairbanks, & Hubbard, 1998). Specifically, the rates of ASPD have been found to be similar in incarcerated males and females (Kokkevi & Stefamos, 1995; Lewis et al., 1982) and in adolescents selected from drug treatment centers (Rounds-Bryant et al., 1998). Thus, although fewer women than men exhibit these disorders in the general population, and although fewer women than men are incarcerated, it appears that many of the women who are in these settings exhibit extreme antisocial behaviors that meet criteria for ASPD.

Returning to the question of predicting base rates in psychopathy, based on the findings from the ASPD/CD literatures, we realize that the question is more complex than we might have expected. Whereas there would be reason to expect gender differences in the base rates of psychopathy given the gender differences in the base rates of delinquent behaviors, it is also possible that the base rates in a prison sample will be equivalent.

Despite the utility of drawing parallels between the extensive ASPD/CD literatures and psychopathy, we should not be too readily guided by these related fields because, although the syndromes overlap in many respects, psychopathy is not the same as ASPD (or CD in adolescents). This distinction between psychopathy and ASPD is often overlooked. Much of the confusion is likely due to the fact that current conceptualizations of psychopathy, particularly the PCL-R, place considerable emphasis on behavioral manifestations of the syndrome. The PCL-R was designed for and has been used primarily with incarcerated (i.e., antisocial) offenders. Factor 2 of the PCL-R, which includes the items “juvenile delinquency,” and “criminal versatility”, is clearly influenced by antisocial behaviors and overlaps the criteria for ASPD in so many ways that individuals often equate them.

However, there is a second component of psychopathy that does not place direct emphasis on antisocial behavior. Captured by Factor 1 of the PCL-R in items such as “callous/lack of empathy,” “grandiose sense of self worth,” and “shallow affect,” this component emphasizes the affective and interpersonal experience of psychopathic individuals, rather than their overt antisocial behaviors.

Although there may be differences in the base rates of criminal and violent behavior of males and females, these behaviors are not synonymous with psychopathy. Rather, they represent just one aspect of the syndrome, which also includes affective and interpersonal components. As a result, when we consider gender differences in the assessment of psychopathy, we will often invoke the literature on ASPD. However, we should not equate the two. Whereas some gender differences in PCL-R–assessed psychopathy may be consistent with gender differences in ASPD, others may not.

TRADITIONAL CONCEPTUALIZATIONS OF THE FEMALE PSYCHOPATH

Although researchers in the field of psychopathy are currently questioning how best to identify the female psychopath, there is good consensus that the psychopathic female does exist. This has been the case since the 1940s, when female psychopaths were included among the case descriptions in Cleckley’s seminal book, The Mask of Sanity.

In his book, Cleckley (1976) provided case descriptions of 15 psychopathic individuals and outlined the 16 core traits of psychopathy that he formulated on the basis of such cases (see Table 1). Two of the 15 case studies presented by Cleckley are those of psychopathic women. Even more important, these two women exhibited the same casual, remorseless, antisocial behaviors and lack of empathy as the male psychopaths. For example, Cleckley (1976) writes of “Roberta,” who, by the time she was 10 years old, was lying without guilt and stealing from family members and shops. As an adolescent, she engaged in truancy, impulsive behavior, and promiscuous sexual behavior. She never seemed remotely concerned with the effect that her behaviors might have on others, yet, at the same time, she never seemed to act in an intentionally malicious way. As her parents told Cleckley, “It’s not that she seems bad or exactly that she means to do wrong . . . . She’s not hard or heartless, but she’s all on the surface . . . she means to stop doing all those terrible things, but she doesn’t mean it enough to matter” (Cleckley, 1976, p. 49).

Cleckley’s (1976) descriptions of psychopathic women highlight the importance of avoiding too narrow a focus on violence and aggression as core features of psychopathy. In part because he does not rely exclusively on specific
behaviors (in contrast to DSM-IV criteria for ASPD), Cleckley's criteria are immune to many gender stereotypes and may be met by women as well as by men. The fact that the prevalence rates of antisocial personality disorder and conduct disorder are lower in females than in males (e.g., APA, 1994; Mulder, Wells, Joyce, & Bushnell, 1994), suggesting that females are less aggressive, ultimately has no bearing on Cleckley's criteria. “Fantastic and uninviting behavior” and “inadequately motivated antisocial behavior” do not specify aggressive behavior. A woman who is sexually promiscuous or tells elaborate lies meets the criteria just as surely as a man who is repeatedly arrested for assault and battery.

Despite the fact that women who met the Cleckley criteria for psychopathy were appearing in clinical settings at least as early as the 1940s, there were few active attempts to study female psychopathy until the 1960s and 1970s, when the number of incarcerated females increased dramatically (Heidensohn, 1994). Even then, the earliest studies of female offenders did not focus on psychopathy per se. Rather, the presence or absence of a psychopathy diagnosis was included only for descriptive purposes, typically within the broader context of a study on female criminality (e.g., Woodside, 1962). Further, such studies often employed different measures of the construct (e.g., clinical descriptions, subscales of the Minnesota Multiphasic Personality Inventory, DSM criteria for antisocial personality), making it difficult to make comparisons across studies or to establish the stability of findings.

When attempting to provide clear answers to the questions surrounding the assessment of the psychopathic female, there are advantages to focusing on a single instrument. For example, focusing on a single instrument allows researchers to draw direct comparisons across numerous studies and samples, as well as to make specific predictions regarding the expectation of gender differences for that particular instrument. It is for this reason that we chose to review data generated in female samples using the PCL-R. Although the PCL-R is not the only measure for assessing psychopathy (e.g., the socialization scale of the California Psychological Inventory [Gough, 1969]; the Self-Report Psychopathy-II [Hare, 1985]; the PCL-Screening Version [Hart, Cox, & Hare, 1995]; the Self-Report Psychopathy Scale [Levenson, Kiehl, & Fitzpatrick, 1995]; the Psychopathic Personality Inventory [Lilienfeld & Andrews, 1996]), the PCL-R provides a reliable and well-validated measure of the psychopathy construct in men. To our knowledge, no other measure of psychopathy equals the PCL-R for empirical demonstrations of reliability and validity. Nor has any other psychopathy instrument been singled out for clinical forensic usage (e.g., risk assessment, treatment planning) in the way that the PCL-R has. Taken together, the evidence for the reliability, validity, and utility of the PCL-R in men provides a powerful rationale for examining the generalizability of this research to female samples.

RESULTS OF STUDIES USING THE PCL-R WITH FEMALE SAMPLES

Reliability

Before we can assess an instrument’s validity, we must first be assured of the measure’s reliability. If there is no diagnostic agreement between studies, across raters, or over time, then predictions made on the basis of a PCL-R rating will not be stable, replicable, or meaningful.

Overall, research supports the reliability of the PCL-R in female samples. Inter-rater reliability in studies with women has consistently been high. A survey of these findings is presented in Table 3. Intraclass correlations range from .87 to .96, indicating good diagnostic agreement. In addition, in the studies where it was assessed, the PCL-R has shown good internal consistency when it has been used in female samples (see Table 3). Coefficient alphas are as high as those typically found in male samples (i.e., .83 to .91; Hare, 1991), which is consistent with the claim that the PCL-R is a relatively homogeneous scale.

Rutherford, Cacciola, Alterman, & McKay (1996) examined the 1-month test–retest reliability of PCL-R categorizations. Test–retest reliability provides an index of how stable a psychopathy classification will be across time. Unfortunately, despite its importance, test–retest reliabil-

<table>
<thead>
<tr>
<th>Study and sample</th>
<th>r</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strachan (1993)</td>
<td>.96</td>
<td>.87</td>
</tr>
<tr>
<td>Total sample</td>
<td>.87</td>
<td>.87</td>
</tr>
<tr>
<td>Rutherford et al. (1996)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total sample</td>
<td>.87</td>
<td>.87</td>
</tr>
</tbody>
</table>

Note: Intraclass correlations and coefficient alphas are presented only from those studies which provided this information.
ity is rarely assessed in studies of psychopathy. Using a score $\geq 25$ to designate psychopathic individuals, Rutherford et al. (1996) found a test–retest reliability of only .40. Although this score is low, it should be noted that Rutherford et al. (1996) could not assess the test–retest reliability using the recommended cut–off score of 30 because none of the participants in this study scored above 29 on the measure. Thus, rather than individuals being classified as “psychopathic” at time one and “nonpsychopathic” 1 month later, the participants in Rutherford et al.’s (1996) study were technically all in the nonpsychopathic range according to current research standards. This use of a non-traditional cut–off score may have artificially lowered the test–retest reliability. Despite these limitations, however, this study represented the first attempt to examine test–retest reliability of the PCL–R in a female sample.

Rutherford et al.’s (1996) inability to use a PCL–R cutting score of 30 due to an insufficient number of women scoring at this level raises the possibility that there may be differences in the number of men and women classified as psychopathic using the traditional PCL–R cut score of 30. A review of the base rate of PCL–R psychopathy in studies with female samples supports this possibility.

### Base Rate

In male prison samples, the PCL–R consistently classifies between 15% and 30% of the inmates as psychopathic (PCL–R $\geq 30$; Salekin, Rogers, Ustad, & Sewell, 1998). In one of the first studies of the PCL–R in a sample of female offenders, Neary (1990) assessed psychopathy in 60 Caucasian and 60 African–American female offenders and found that the distribution of PCL–R scores was dissimilar to that found in men, with only 11% of the sample meeting the PCL–R criteria for psychopathy (PCL–R $\geq 30$). Louks (1995) also found a base rate of 11% in her sample of incarcerated females. In a study of 103 female inmates, Salekin, Rogers, and Sewell (1997) found that 16% of their sample scored 30 or above on the PCL–R. Higher rates of psychopathy were found by Tien, Lamb, Bond, Gillstrom, and Paris (1993), who found that 23% of their sample of 74 female inmates met PCL–R criteria for psychopathy, a base rate closer to that found in male prison samples. Thus, although there are women who meet the PCL–R criteria for psychopathy, on the basis of these preliminary studies, the prevalence of psychopathy among incarcerated female offenders appears to be lower than among male offenders. See Table 4 for a survey of base rates across studies.

Support for the existence of comparable numbers of male and female psychopaths was provided by Strachan (1993), who examined PCL–R assessed psychopathy in a sample of 75 female Canadian inmates. Using the recommended cut–off score of 30 to identify individuals as psychopathic, Strachan (1993) found that 31% of her sample was classified as psychopathic. This prevalence is as high as the rates typically assessed among male offenders. The author attributed this finding to the fact that 35 of the 75 women included in her sample were incarcerated at the maximum security level (Strachan, 1993).

Rutherford et al.’s (1996) examined PCL–R scores in a nonoffender sample of 58 female methadone patients. The mean PCL–R score was 13.8, which is considerably lower than the mean score in previous female samples. This finding could be attributed to the different compositions of the samples. Whereas Neary (1990), Tien et al. (1993), Strachan (1993), and Salekin et al. (1997) assessed incarcerated offenders, Rutherford et al.’s (1996) sample consisted of nonincarcerated women in a clinical, as opposed to forensic, setting. However, it is interesting to note that the mean PCL–R score in this sample was significantly lower than the mean PCL–R score in a comparable sample of male methadone patients (Alterman, Cacciola, & Rutherford, 1993). Similarly, Weiler & Widom (1996) conducted a within–sample gender comparison of PCL–R scores in a noninstitutionalized sample that included both abused or neglected young adults and a matched control group and found that men’s PCL–R scores were higher than women’s ($M = 8.8$ vs. 4.6, respectively). Thus it appears that men tend to show higher PCL–R scores than women in incarcerated, clinical, and community samples.

The disparities in the distributions of PCL–R scores both across female samples and between male and female

---

### Table 4. Base rates of PCL–R assessed psychopathy in female samples

<table>
<thead>
<tr>
<th>Study and sample</th>
<th>Base rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rutherford et al. (1996) Total sample</td>
<td>0%</td>
</tr>
<tr>
<td>Neary (1990) Total sample</td>
<td>11%</td>
</tr>
<tr>
<td>Louks (1995) Total sample</td>
<td>11%</td>
</tr>
<tr>
<td>Salekin et al. (1997) Total sample</td>
<td>16%</td>
</tr>
<tr>
<td>Tien et al. (1993) Total sample</td>
<td>23%</td>
</tr>
<tr>
<td>Strachan (1993) Total sample</td>
<td>31%</td>
</tr>
</tbody>
</table>
samples highlight the fact that base rates may vary widely as a result of sample composition. This is analogous to the finding in the ASPD literature that gender discrepancies in base rates may exist in some samples but not in others. For example, it has been suggested that Rutherford et al. (1996) failed to find any psychopathic women due to the fact that they were using a clinical as opposed to a forensic population. Strachan (1993) argued that her unusually high base rate was due to the proportion of maximum security inmates included in her sample. Even differences in the prison policies between nations have been implicated in explanations for the higher base rates of psychopathy in Canadian prison samples as compared to American prison samples (D. Mailloux, personal communication, August 16, 1998).

When considering the disparate base rates across female samples, it is important to note that, if the base rates generated from the unusual samples used by Strachan (1993) and Rutherford et al. (1996) are excluded, we find that the base rate of psychopathy in incarcerated women ranges from 11% to 23%. This range is not very different from the range of 15–30% found in men. Thus, although the prevalence of psychopathy in women is generally lower than in men, there is some overlap in the base rates of the syndrome across male and female samples.

If the base rate of psychopathy within prison populations, which have long been believed to provide an “enriched” sample for the study of psychopathy, is lower for women than for men, does this also mean that the prevalence in the general population will be lower in females than in males? This is not necessarily the case. It may be that female psychopathic individuals’ antisocial behaviors are less overt and less likely to result in incarceration than those of male psychopathic individuals. For example, a female psychopathic individual who neglects her young children may do so without drawing the attention of authorities. Another possibility is that the psychopathic female uses manipulation to coerce others to commit antisocial acts that she orchestrates. Although these are only speculations, both of these scenarios provide examples of how a female psychopathic individual in the community could behave in antisocial ways and yet avoid incarceration.

A second possibility is that gender bias in sentencing leads to fewer incarcerated female psychopathic individuals than male psychopathic individuals. This position is supported by data on adolescents showing that female adolescents are more likely to have their cases handled informally by the justice system than male adolescents and that they are less likely than their male counterparts to be remanded to adult courts (Hoyt & Scherer, 1998).

A related issue highlighted by the existing research that is consistent with debate surrounding the diagnostic criteria for CD and ASPD is the possibility that the base rates of psychopathy appear lower in women because the PCL-R is not adequately assessing psychopathy in women. If psychopathy were manifested differently in males and females, then it is possible that the items composing the PCL-R are not sufficient for identifying female psychopaths. Although most items appear to be applicable to both men and women on the basis of their gender-neutral wording, it may be that the existing items are not sampling relevant domains. For example, quality of parenting may represent a domain wherein psychopathic attributes are manifested by women. For example, research has suggested that psychopathic women are more likely to surrender their children to adoption (Strachan, 1993). It does not follow from this finding, however, that surrendering children for adoption should be made a criteria for female psychopathy. Rather, the finding suggests first that the motherhood of female psychopaths should be further investigated, and second, that it might be useful to examine child-rearing practices in greater depth and/or place greater emphasis on them when assessing psychopathy in women, particularly with regard to how these behaviors may inform other aspects of PCL-R psychopathy (e.g., “irresponsibility”).

It may also be that existing items do not have the same sensitivity in women as they do in men. The sensitivity of each item is based, in part, on the range of behaviors available to exemplify the item and the representativeness of those examples. If women engage in fewer exemplar behaviors, then we would not expect the items to be as sensitive. For example, when raters are making a determination of “callousness,” they may find that an extensive history of physical violence by the individual will contribute to higher scores on this item. In the absence of physical aggression, there are fewer clear examples of the characteristic to use as evidence for a high rating on the item.

Another difficulty with current PCL-R items may be the emphasis on formal charges or convictions. For example, if the rates of female juvenile delinquency are lower in women than in men (e.g., Shannon, 1981; Tracy, Wolfgang, & Figlio, 1985), then using the existence of formal criminal charges as a juvenile as one of the criteria
for defining psychopathy would necessarily lead to fewer women than men being classified as psychopathic.

Each of the possibilities described above provides an example of potentially problematic items. However, it is not necessary to rely on subjective opinions when exploring potential item bias in an instrument. Rather, empirical approaches can be used to examine the contributions and function of individual items to the classification of psychopathic women. Recommended methods include the use of factor analysis and item response theory analyses to compare the structure of the instrument and the functions of the items in women to those in men. Thus, we next consider the research which has examined the structure of the PCL–R in female samples.

Factor Structure
In one of the few studies to address the factor structure of the PCL–R in female offenders, Salekin et al. (1997) conducted an exploratory factor analysis on a sample of 103 participants to compare the factor structure of the PCL–R in a female sample with the two-factor structure reliably found in male samples. Interestingly, the authors noted that the factor structure in their sample was not the same as the factor structure typically found with men. Although the two factors appeared to be somewhat similar to the two factors reported by Harpur, Hare, and Hakstian (1989), one being marked by interpersonal traits, and the other by socially deviant behaviors, the individual PCL–R items did not load on these factors in the same way as they do in male samples. Three of the items (poor behavioral controls, impulsivity, and lack of realistic, long-term goals) cross-loaded, and three of the items (failure to accept responsibility, many short-term marital relationships, and revocation of conditional release) failed to load > .4 on any factor.

On the basis of their data, Salekin et al. (1997) concluded that in women Factor 1 is “characterized by lack of empathy or guilt, interpersonal deception, proneness to boredom and sensation seeking,” whereas Factor 2 is “characterized by early behavioral problems, promiscuity, and adult antisocial behavior” (p. 582). This difference in factor structure may indicate important differences in the latent construct that the measure is assessing. As Ben-Porath (1990) notes, when factor structures differ between cultural or racial groups, “the investigator is alerted that a qualitative change has occurred in the instrument” (p. 33). As a result, it may be inappropriate to assume that the correlates of the psychopathy construct in females will be identical to those which represent the construct in males.

Salekin et al.'s (1997) finding should be accepted cautiously, however, for two reasons. First, their sample of only 103 participants is too small to conduct adequate exploratory factor analysis on the PCL–R. The most recent view is that an exploratory factor analysis for an instrument of the PCL–R’s length and structure requires a sample size of at least 200 participants (MacCallum, Widaman, Zhang, & Hong, 1999). Second, Salekin et al. (1997) included both Caucasian and African-American women in their analyses. Studies examining the factor structure of the PCL–R across race in male samples have found differences between Caucasian and African-American samples (Kosson, Smith, & Newman, 1990; Lorenz, Smith, Bolt, Schmitt, & Newman, 1999). Just more than one half of Salekin et al.’s (1997) sample of 103 participants was composed of Caucasian women. Thus, the failure to replicate the factor structure typically observed in Caucasian, male samples is qualified by the small sample size and the failure to consider race.

There is a large gap in the literature on the applicability of the PCL–R to female samples in terms of the structure and item function of the instrument. The single published study to address the factor structure of the instrument (Salekin et al., 1997) is inadequate due to a small sample and a failure to disaggregate the sample by race. As a result, there are clear implications for future research. First, researchers examining PCL–R psychopathy in females need to assess an adequate number of participants so that proper EFAs can be conducted. Specifically, at least 200 participants of a single race should be assessed. Second, researchers working with female samples should begin using item response theory methodologies to investigate individual item functioning.

In item response theory models, the performance of each item of the instrument can be described on the basis of two parameters: parameter a is relevant to how well the item discriminates on the latent trait, and parameter b concerns at what level of the latent trait the item discriminates (Cooke & Michie, 1997). For example, some characteristics of the disorder may not be apparent until the individual is extreme on the latent trait, whereas other characteristics are apparent at much lower levels of the latent trait. Item response theory methods have already been used to examine differences between Scottish and North American male samples (Cooke & Michie, 1997) and between Caucasian and African-American male
samples (Lorenz et al., 1999). It would be a useful and natural extension to employ item response theory methods in comparisons involving male and female samples.

In the absence of adequate data examining the structure and item function of the PCL-R in women, we must direct our attention to other methods of evaluating the PCL-R’s validity in women. Specifically, we can examine whether the PCL-R is tapping a similar construct in women as in men, on the basis of the correlates of this construct. If the construct is dissimilar in men and women, then we cannot generalize findings from studies with male offenders to populations of female offenders. Further, the clinical utility of the instrument is based on its validity. If there is insufficient evidence for the validity of the instrument in women, then we should not make the same clinical predictions using the instrument with women as we do with men.

Substantive Validity
Because the majority of studies that have examined the reliability of the PCL-R in female samples have also examined the convergent, discriminant, and/or predictive validity of the PCL in these samples, we can review the evidence for the validity of the PCL-R in female samples. Further, because these validity measures have included both overt behaviors (e.g., criminal convictions, alcohol abuse) and personality characteristics (e.g., empathy, narcissism, anxiety), we can draw a distinction between those measures that primarily relate to the antisocial behavioral component of psychopathy and those that primarily relate to the personality characteristics or attitudes associated with the disorder.

One common approach to examining the validity of the PCL-R has been to relate PCL-R scores to scores obtained on various self-report measures of psychopathy. Results of these analyses have shown that PCL-R scores are significantly associated with scores on Hare’s (1985) Self-Report Psychopathy-II (SRP-II) Scale (Rutherford et al., 1996; Strachan, 1993). Such findings demonstrate that the PCL-R is related to other measures of the psychopathy construct. Further, the agreement between the interview measure and self-report instruments shows that PCL-R raters are conceptualizing their participants in ways similar to how the participants are viewing themselves.

In addition to self-report psychopathy measures, studies of the PCL-R in female offenders have included measures chosen to tap the unsocialized, antisocial behavior that is also characteristic of the disorder (see Cleckley criteria in Table 1). These have included self-report measures such as the socialization (SO) subscale from the California Psychological Inventory (Gough, 1969), as well as actual behaviors, represented by diagnoses of ASPD and substance abuse according to DSM-III-R criteria (APA, 1987), and records of criminal acts and recidivism.

Despite gender differences in the general rates of antisocial behavior, research provides moderate support for the proposition that women classified as psychopathic according to the PCL-R resemble male psychopathic individuals behaviorally. For example, Strachan (1993) showed that PCL-R scores were significantly associated with lower scores on the SO scale \((r = -0.48, p < .01)\), although Rutherford et al. (1996) failed to show a significant relationship between the two measures.

Examination of the relationship between PCL-R scores and substance abuse have also been somewhat inconsistent. Although Rutherford et al. (1996) found no significant relationship between PCL-R total scores and either a diagnosis of alcohol dependence or the number of substance use disorders an individual had experienced, they did find a significant association between PCL-R Factor 2 scores and the number of substance use disorders an individual had experienced. This is consistent with findings in males that substance abuse appears to be associated primarily with Factor 2 (e.g. Hart & Hare, 1989; Smith & Newman, 1990).

As in studies of male samples, the relationship between PCL-R scores and criminal behavior has been a robust in female samples. Strachan (1993) found that psychopathic individuals were significantly more likely than nonpsychopaths to have a prior conviction \((\chi^2 = 9.01, p < .005)\), and to have been convicted of one or more violent offenses \((\chi^2 = 4.51, p < .04)\). Further, psychopathic individuals were more likely to have engaged in prostitution \((t = 3.22, p < .01)\). Rutherford et al. (1996) found a significant association between PCL-R scores and the number of previous arrests \((r = .42, p < .001)\). Finally, a relationship between PCL-R scores and criminal behavior has been noted in community samples. Weiler and Widom (1996) showed that, in their noninstitutionalized, community sample, PCL-R scores significantly predicted violence as measured both by the number of arrests for violent crime and self-reported violence.

The consistent relationship between PCL-R scores and criminal behavior (Rutherford et al., 1996; Strachan, 1993; Weiler & Widom, 1996) provides evidence that the instrument may be as sensitive to differences in criminality
in female samples as in male samples. However, because PCL-R ratings will be influenced by past criminal activity, it is also necessary to examine the PCL-R’s ability to predict future criminality. Past research with male offenders has supported the predictive validity of the PCL-R in males by showing that PCL-R–assessed psychopathy is a good predictor of criminal recidivism (e.g., Harris, Rice, & Cormier, 1991; Hemphill et al., 1998; Salekin, Rogers, & Sewell, 1996; Serin & Amos, 1995; Serin et al., 1990).

In an attempt to examine the predictive validity of the PCL-R in a female sample, Salekin et al. (1998) examined recidivism data from 78 female inmates. The results of the study revealed a somewhat dissimilar relationship between PCL-R scores and recidivism in females from that observed in males. First, whereas studies with males have found a mean reoffending rate of 62.6% among individuals classified as psychopathic using the PCL-R (Salekin et al., 1998), Salekin et al.’s (1998) study of recidivism in women found that the rate for psychopathic females in their sample was only 50%. Further, the study showed that the classification accuracy for psychopathy as a recidivism predictor was “moderate to poor” (Salekin et al., 1998). Using the PCL-R cut-off score of 30, approximately nine-tenths of the women who recidivated had not been classified as psychopathic. In addition, 9% of the individuals who did not recidivate were classified as psychopathic (Salekin et al., 1998). Thus, using the PCL-R to predict recidivism in this sample resulted in high rates of false negatives and a smaller percentage of false positives.

In a related study, Salekin et al. (1997) found that personality-based measures of psychopathy, including the PCL-R, were unrelated to subsequent ratings of violent behavior, verbal aggression, noncompliant behavior, and overall dangerousness as scored by correctional officers who had contact with the female offenders. Thus, it appears that PCL-R ratings of women have not yet been shown to adequately predict institutional adjustment.

These findings highlight the difference between predicting and postdicting criminal behavior. Whereas PCL-R ratings in studies of females are related to past criminal behavior (Rutherford et al., 1996; Strachan, 1993; Weiler & Widom, 1996), Salekin et al.’s (1998) finding suggests that the PCL-R may not be as good a predictor of future criminal behavior in women. Thus, when PCL-R scores are related to previous criminality, it is unclear if this is because the PCL-R is tapping the capacity for future criminal behavior or because previous criminal behavior is greatly influencing PCL-R scores.

Overall, there is evidence for similarities in the behavioral correlates of PCL-R–assessed psychopathy in men and women. This is interesting in light of concerns that gender differences in the rates and expression of antisocial behavior might lead to gender differences in the behavioral expressions of psychopathy. In spite of these concerns, it appears that there are female offenders whose past antisocial behaviors are tapped adequately by the PCL-R. What is now required are additional studies of the predictive capabilities of the PCL-R in women to determine if PCL-R psychopathy is associated with the same types of antisocial behaviors as it is in men and to determine if the instrument has the same forensic utility in women as it does in men. Studies should examine the relationship between PCL-R scores and subsequent criminal recidivism, institutional adjustment, and treatment outcome.

Despite the importance of the affective and interpersonal components of psychopathy, more research has examined the behavioral correlates of psychopathy in women than these other characteristics. The studies that have been conducted have used personality measures selected to tap the glib, callous, grandiose, and unempathic characteristics that help to define the psychopathy construct. Measures have included the Interpersonal Reactivity Index (Wiggins, 1979), the Narcissistic Personality Inventory (Raskin & Hall, 1979), the Machiavellianism–IV scale (Christie & Geis, 1970), the Eysenck Personality Questionnaire–Revised (Eysenck & Eysenck, 1975), and various measures of trait anxiety, including the State–Trait Anxiety Inventory and the Welsh Anxiety Scale (Welsh, 1956).

Results have provided some support for the position that the PCL-R is capturing the various personality characteristics in women that have previously been associated with psychopathy in males. Strachan (1993) showed that PCL-R scores were significantly negatively associated with the Perspective taking subscale of the Interpersonal Reactivity Index ($r = -.31, p < .05$). Similarly, Rutherford et al. (1996) found that PCL-R scores were significantly associated with lower scores on the perspective taking scale ($r = -.57, p < .001$) and the empathic concern scale of the Interpersonal Reactivity Index ($r = -.57, p < .001$).

Rutherford et al. (1996) also included the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975)
and found that as expected, PCL-R scores were significantly associated with psychoticism ($r = .48$, $p < .001$). PCL-R scores were not associated with scores on the extraversion, neuroticism, or lie scales. Finally, the authors found no significant relationship between PCL-R scores and scores on the Machiavellianism-IV scale (Christie & Geis, 1970).

Assessments of anxiety have shown that, as in male samples, the PCL-R appears to be independent of scores on anxiety scales. For example, Strachan (1993) showed that PCL-R scores were independent of scores on the State-Trait Anxiety Inventory, and Rutherford et al. (1996) showed that PCL-R scores were independent of the presence of a diagnosis for an anxiety disorder, as well as scores on the neuroticism subscale of the Eysenck Personality Questionnaire. PCL-R scores have also been demonstrated to be independent of depression diagnoses (Rutherford et al., 1996).

Although the results are promising, there are surprisingly few data available on the affective and interpersonal characteristics of psychopathic women. Previous studies have failed to examine important relationships such as the association between PCL-R scores and IQ scores. Cleckley (1976) specifically stated that the psychopathic individual’s behavior was not the result of low intelligence. Thus, it is necessary to show that increasing PCL-R scores do not merely reflect decreasing IQ scores. However, researchers have failed to include even brief measures of IQ in their studies of the PCL-R in women.

Perhaps even more striking than the absence of additional personality measures in studies of the PCL-R in women is the absence of research examining the important laboratory correlates of the syndrome. Research using the PCL-R in males has provided laboratory evidence for cognitive and emotional processing deficits in psychopathic individuals that are theoretically linked to deficient self-regulation and therefore relevant to understanding the etiology of psychopathy. Psychopathic individuals are deficient in their ability to learn from punishment, as demonstrated by their poor passive avoidance learning (e.g., Lykken, 1957; Newman & Kosson, 1986; Newman, Patterson, Howland, & Nichols, 1990; Newman & Schmitt, 1998; Thornquist & Zuckerman, 1995). Psychopathic individuals also have difficulty shifting their attention to process secondary information when they are engaged in a dominant response. This deficit in response modulation has been shown in a variety of laboratory tasks, including tasks using secondary punishment cues (e.g., passive avoidance tasks), and motivationally neutral secondary cues (e.g., Newman, Schmitt, & Voss, 1997; Schmitt & Newman, 1999). Finally, the psychopathic individual’s deficiency in processing emotion cues has also been demonstrated using a variety of tasks, ranging from lexical decision tasks involving emotion words (e.g., Lorenz & Newman, 2000; Williamson et al., 1991) to measures of psychophysiological responses to threat (e.g., Arnett, Howland, Smith, & Newman, 1993; Ogloff & Wong, 1990; Patrick, Bradley, & Lang, 1993).

The failure to extend these theoretically driven examinations of psychopathy to female samples represents a major limitation of the field at this time. Current evidence for the validity of the PCL-R in women focuses only on the description of the psychopathic woman’s attitudes and behaviors. What is required are examinations of the processes theorized to underlie these behaviors and attitudes. Without studying the processes that are purported to cause the syndrome, we will never know if gender differences in the descriptive correlates of the syndrome are the result of differences in how the processes are expressed or differences in these fundamental causal mechanisms themselves.

Overall, there is some evidence for the validity of the PCL-R in female samples. For example, there are a number of findings establishing an association between women’s PCL-R scores and scores on measures assessing a variety of theoretically related constructs (e.g., past criminal behavior, poor perspective taking). Although promising, these findings are limited by the fact that many of the measures which have been used to examine the validity of psychopathy in women have been used in only one or two studies, thereby preventing us from determining the replicability of these associations. Further, researchers need to focus more specifically on the discriminant validity of the PCL-R. Although investigators have consistently assessed anxiety and depression and have found them to be independent of PCL-R scores, measures of other forms of psychopathology and measures of intelligence have been neglected.

**IMPLICATIONS FOR CLINICAL USAGE**

The overlapping goals of this paper were to determine if the PCL-R provides reliable and valid assessments of the psychopathy construct in women and to present information that would assist mental health professionals in evalu-
ating the utility of applying the PCL-R to women. First, it was necessary to assess the extent to which the PCL-R, when used with women, meets Hare’s (1991) standards for clinical usage. The majority of empirical evidence shows that the PCL-R is a reliable instrument when used in female samples. However, the evidence for the validity of the PCL-R is less clear. Although the research looks promising, we must determine if this evidence is sufficient for considering the measure validated, and therefore applicable, within female samples. Regardless of how a clinician is using the instrument (e.g., traditional psychopathy groupings, dimensional usage), if the instrument is valid, then researchers and clinicians should be able to make specific predictions about the description, etiology, treatment, and prognosis of psychopathy.

If the PCL-R is a valid measure of psychopathy in women, then PCL-R scores should provide clinicians with information relevant to the description or symptomatology of psychopathy. Knowing an individual’s PCL-R classification or level should enable clinicians and researchers to predict the existence of those beliefs and behaviors that are currently used to define psychopathy. For example, if clinicians were to classify a man as psychopathic using the PCL-R, those clinicians could make a number of predictions based upon extensive research on male psychopaths. They could predict, among other things, that the man would be likely to show high levels of future antisocial behaviors, be noncompliant in treatment, and that he would show poor perspective taking and empathy.

The important question for our purposes is, could those clinicians make these same predictions if they had just used the PCL-R to classify a woman as a psychopath? We know that PCL-R–assessed psychopathy is associated with poor perspective taking and decreased empathy (Rutherford et al., 1996; Strachan, 1993). Further, there appears to be an association between PCL-R scores and higher levels of past criminal behavior in women (Rutherford et al., 1996), although not necessarily with institutional aggression (Salekin et al., 1997) or increased risk for recidivism (Salekin et al., 1998). Thus, whereas they might be able to postdict that their client has engaged in aggressive behavior in the past, there is no evidence that they would be able to predict her future criminal behavior. When considering the likelihood of treatment noncompliance, we find there is little evidence for an association between PCL-R scores in females and unsuitability for treatment. In fact, in the one study that considered treatment compliance, Salekin et al. (1997) found no relation between PCL-R scores and scores on the treatment rejection scale of the Personality Assessment Inventory (Morey, 1991).

The research in many areas is limited. The majority of research on females has focused on the symptomatology of psychopathy at the expense of studying etiology, prognosis, or treatment. Further, even the studies of symptomatology have failed to provide clear and consistent answers. This might in part be due to the lower base rate of psychopathy in women, which makes it more difficult to study psychopathic women than men simply because there are fewer of them. Whatever the cause, the result is that, in contrast to findings with males, there is little evidence regarding the presence of etiologically relevant cognitive-behavioral deficiencies, psychopathic females’ propensity to recidivate, or their ability to benefit from treatment relative to nonpsychopathic women.

When determining whether to use a PCL-R assessment of a female, researchers or clinicians must ask themselves, what is their reason for choosing to conduct that assessment? Throughout this paper, various uses for the PCL-R have been mentioned, including predicting criminal recidivism, predicting institutional violence, and planning and implementing treatments and interventions. If clinicians were using the PCL-R for the sole purpose of predicting specific outcomes for any particular woman in these areas, they would be doing so without empirical evidence of the predictive power of the PCL-R in such domains. That evidence simply has not been collected yet. Thus, we caution against the premature use of the instrument to make important decisions based upon PCL-R ratings (e.g., inclusion of scores in a risk assessment battery, use of scores to determine potential benefit from treatment). Conversely, the clinician or researcher who uses the PCL-R to assess psychopathy with the intention of clarifying the relationships which exist or discovering new ones should do so. This is the work that is needed to understand and refine the assessment of psychopathy in women so that this assessment will have greater utility in the future. This work cannot be conducted if the PCL-R is abandoned.

The reality is that far more data need to be collected on the relationship between PCL-R–assessed psychopathy and women’s criminal behavior, parenting behavior, substance use, and treatment outcomes before the instru-
ment can be considered a useful tool for predicting outcomes in these areas. Further, at this time there is insufficient evidence to guide alteration of the PCL-R. Thus, we should continue to use the PCL-R in its current form until we have sufficient theoretical and/or empirical rationales for changing it. It is our belief that such rationales will follow from research that directly addresses the etiology of psychopathy in women. Understanding the etiological processes underlying psychopathy in women will not only guide decisions regarding the PCL-R and its use in these samples, but will clarify the factors that contribute to the development of psychopathy and will highlight those variables that must be altered to prevent this serious disorder.

**ACKNOWLEDGMENT**
Preparation of the manuscript was supported by a research grant from the National Institute of Mental Health. We thank William Schmitt, Chad Brinkley, Amanda Lorenz, and Donal MacCoon for their comments on early drafts of this manuscript.

**NOTE**
1. In addition to using information obtained from interviews to make PCL-R ratings, Weiler & Widom (1996) also used information obtained through available criminal records and case files to rate their participants.

**REFERENCES**
Two separate diagnoses or two aspects of the same psychopathy? Comprehensive Psychiatry, 38, 237–242.


Received February 14, 2000; revised June 2, 2000; accepted June 16, 2000.