

The Revised Psychopathy Checklist: Reliability and Factor Structure

Robert D. Hare, Timothy J. Harpur, A. R. Hakstian,
Adelle E. Forth and Stephen D. Hart
University of British Columbia
Vancouver, British Columbia, Canada

Joseph P. Newman
University of Wisconsin

The revised Psychopathy Checklist (PCL) is a 20-item scale scored from interview and file information. Analyses of data from 5 prison samples ($N = 925$) and 3 forensic psychiatric samples ($N = 356$) indicate that the revised PCL resembles its 22-item predecessor in all important respects. It has excellent psychometric properties, and it measures 2 correlated factors that were cross-validated both within and between samples. Correlations between the original PCL and the revised version approached unity for both the factors and the full scale. We conclude that the revised PCL measures the same construct as the original and that the PCL is a reliable and valid instrument for the assessment of psychopathy in male forensic populations.

During the last decade, we have devoted considerable effort to the development of an assessment procedure for criminal psychopathy that has acceptable psychometric properties and that is closely tied to traditional clinical conceptions of psychopathy. The result is the Psychopathy Checklist (PCL).

The original PCL consists of 22 personality and behavioral items, which are completed on the basis of interview and file information (Hare, 1980). There is a considerable amount of published and unpublished evidence attesting to the reliability and validity of the PCL.¹ The PCL has been shown to measure two correlated factors (Harpur, Hakstian, & Hare, 1988). There are differences in external correlates of the factors, and we have argued that measurement of both factors is important for a comprehensive assessment of psychopathy (Harpur & Hare, 1989; Harpur, Hare, & Hakstian, 1989).

The revised Psychopathy Checklist (PCL-R; Hare, 1985) differs from the PCL in several ways. Two PCL items, *previous diagnosis as a psychopath or similar* and *drug or alcohol not direct cause of antisocial behavior*, have been omitted because they had relatively low correlations with the total score and were difficult to score. The scoring of some items was made more stringent by requiring more extreme instances of deviant behavior for a score of 2. For a few items, the scoring criteria were modified somewhat.

It is important for the continuity of research carried out using the PCL to demonstrate empirically that the revised version

does not differ substantially from the original. Even minor changes in scoring criteria can result in substantive changes in the meaning of an item. The purpose of this article is to demonstrate that the changes introduced into the PCL-R have not altered its psychometric properties and that the body of research accumulated using the PCL can be generalized to the revised measure.

Method

We obtained PCL-R item scores for five samples of male prison inmates (Samples 1-5; $n = 925$) and three samples of male forensic psychiatric patients (Samples 6-8; $n = 356$). These samples were completely independent of those reported by Harpur et al. (1988). All of the inmates in Sample 3 and at least 85% of those in the other samples were White; most of the other inmates were Native American. Subjects' ages ranged from 16 to 69 years, but the majority were under age 40 (see Harpur & Hare, 1989, for a detailed breakdown of the ages of four of the current samples). Because inmates suffering from serious mental disorders are usually transferred to forensic hospitals, the prison samples generally did not include any psychotic patients. Subjects in the forensic psychiatric samples suffered from a variety of mental disorders, including psychotic disorders (see Hart & Hare, 1989).

1. Sample 1 consisted of 241 inmates of a federal medium-security institution in British Columbia, Canada, all of whom were serving sentences of 2 years or longer. These subjects had volunteered to take part in several research projects. Two ratings were available for 176 of the inmates.

2. Subjects in Sample 2 were 122 inmates at a provincial prison in British Columbia, all serving sentences of less than 2 years. They had volunteered to participate in several laboratory and paper-and-pencil tests. Two ratings were available for 37 of the inmates.

3. Sample 3 consisted of 369 inmates at a minimum-security institu-

Preparation of this report was supported by Grant MT-4511 from the Medical Research Council of Canada to Robert D. Hare and by Grant MH37711 from the National Institute of Mental Health to Joseph P. Newman.

Correspondence concerning this article and requests for an extended report should be addressed to Robert D. Hare, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y7.

¹ A mimeographed manual for this revised PCL is available on request, and a more formal manual is in preparation. A current bibliography of research using the PCL will be included in the manual.

tion in Wisconsin. They had volunteered to participate in several experiments conducted by Joseph Newman and his colleagues. Two ratings were available for 72 of the inmates.

4. Sample 4 consisted of 106 inmates who were assessed, using a French translation of the PCL (Côté & Hodgins, 1989), just prior to conditional release from minimum-, medium-, or maximum-security federal prisons in Quebec. Two ratings were available for 70 of the inmates.

5. Sample 5 consisted of 87 inmates in a medium-security prison in Kingston, Ontario. They had volunteered to participate in research projects conducted by Ralph Serin. Only one rating was available for each inmate.

6. Sample 6 consisted of 80 patients who had been remanded to a forensic hospital in British Columbia for inpatient assessment of fitness to stand trial (see Hart & Hare, 1989). Two raters completed the PCL-R for each patient.

7. Sample 7 consisted of 165 patients in an intensive therapeutic program for mentally disordered offenders in a forensic psychiatric unit of the Mental Health Center at Penetanguishene, Ontario, Canada (Rice & Harris, 1989). Only one rating was available for each patient.

8. Sample 8 consisted of 111 patients at the Regional Psychiatric Center (RPC) in Saskatchewan, Canada. The RPC is a forensic psychiatric hospital that accepts inmates from federal prisons. The ratings were collected by Stephen Wong. Only one rating was available for each inmate. We did not have access to item scores for this sample.

Each PCL-R item was scored on a 3-point scale (0, 1, 2) according to the degree to which it applied to the individual. Occasionally, it was not possible to complete an item with confidence, in which case the item was omitted, and the total score was prorated to a 20-item scale. For computation of coefficient alpha, the value 1 was assigned to omitted items. Different raters were involved in assessments made at different sites. They included clinical psychologists, graduate students, and research assistants. Where possible, a second rater completed the PCL-R independently after observing the original interview (which was either live or recorded on videotape) and reviewing the file information. When available, these ratings were averaged together.

For the five prison samples and for Sample 8, items in the PCL were scored by integrating interview and file information. However, some of the patients in Sample 6 and all of those in Sample 7 could not be interviewed because they were no longer at the institution; for these patients, PCL-R ratings were made on the basis of extensive correctional and hospital files (see Wong, 1988).

We performed factor analyses on Samples 1, 2, 3, 4, and 7. Samples 5 and 6 were omitted because of small sample sizes, and item scores for Sample 8 were not available. Two subjects from Sample 7, for whom more than four PCL items could not be scored because of missing data, were dropped for these analyses. Correlation matrices were calculated after replacing missing data by item means for that sample.

We followed the procedures used by Harpur et al. (1988), to which the reader is referred for details. We used split-half cross-validation (Everett, 1983) on Samples 1 and 3. We next compared congruence coefficients for the best two-, three-, and four-factor oblique solutions from the five data sets independently. A single invariant factor pattern was subsequently obtained for Samples 1, 3, 4, and 7, using the Method I procedure described by Meredith (1964). Finally, the psychometric properties of scales derived from the factors were examined.

Results and Discussion

Full-Scale PCL

The items in the PCL-R are listed in Table 1, along with the mean item-total correlation for Samples 1-7 and the common primary-factor pattern for the four samples included in the

Table 1

Revised Psychopathy Checklist Items, Their Mean Item-Total Correlations, and the Common Primary-Factor Pattern

Item	Mean item-total correlation ^a	Factor ^b	
		1	2
1. Glibness/superficial charm	.50	<u>.86</u>	-.25
2. Grandiose sense of self-worth	.52	<u>.76</u>	-.16
3. Need for stimulation	.58	<u>.09</u>	<u>.56</u>
4. Pathological lying	.53	<u>.62</u>	<u>.03</u>
5. Conning/manipulative	.58	<u>.59</u>	.10
6. Lack of remorse or guilt	.53	<u>.53</u>	.11
7. Shallow affect	.58	<u>.57</u>	.10
8. Callous/lack of empathy	.64	<u>.53</u>	.22
9. Parasitic lifestyle	.44	-.00	<u>.56</u>
10. Poor behavioral controls	.50	.14	<u>.44</u>
11. Promiscuous sexual behavior	.37	.35	<u>.08</u>
12. Early behavior problems	.46	-.01	<u>.56</u>
13. Lack of realistic goals	.50	.10	<u>.56</u>
14. Impulsivity	.52	.01	<u>.66</u>
15. Irresponsibility	.53	.16	<u>.51</u>
16. Failure to accept responsibility	.38	<u>.47</u>	<u>.02</u>
17. Many short-term relationships	.27	.18	.18
18. Juvenile delinquency	.32	-.18	<u>.59</u>
19. Revocation of conditional release	.35	-.00	<u>.44</u>
20. Criminal versatility	.43	.15	<u>.33</u>

^a Mean of Fisher-Z transformed, corrected item-total correlations for Samples 1-7, weighted by sample size.

^b Coefficients > 0.4 have been underlined. Coefficients for the common solution are scaled so that the column sums of squares are equal to those of the average of the four sample solutions.

pooled-factor solution. The means, medians, and standard deviations of the PCL-R total scores were highly similar for both the prison and forensic psychiatric samples. Means ranged from 20.1 to 23.9 ($M = 22.1$), medians from 20.0 to 25.0 ($M = 22.8$), and standard deviations ranged from 6.7 to 9.0 ($M = 7.95$). The distribution of scores within each sample was approximately normal. Although several different types of institutions and many different raters were involved, the characteristics of each sample within each type of setting were very similar to one another.

To assess interrater reliability, intraclass correlation coefficients (ICCs; Shrout & Fleiss, 1979) were calculated for subsamples of subjects rated by more than one rater (see sample descriptions for N s). They ranged from .78 to .94 ($M = .86$) for a single rating and from .87 to .97 ($M = .93$) for the average of two ratings (computations based on the Spearman-Brown Prophecy Formula).² Internal consistency was assessed by Cronbach's coefficient alpha, which ranged from .85 to .89 ($M = .88$), and by the mean interitem correlation, which ranged from .23 to .30 ($M = .27$). These results indicate that the PCL-R, like the PCL, can be considered a homogeneous, unidimensional scale (Harpur et al., 1989). There were no apparent differences in reliability coefficients obtained in prison versus forensic psychiatric samples.

² Whenever possible, assessments should be based on the average of at least two independent ratings.

Factor Structure

Factor comparability coefficients were calculated for Samples 1 and 3, with from two to five factors extracted. Only the two-factor solutions produced factors that cross-validated within each sample. The results for the PCL-R were very similar to those presented by Harpur et al. (1988) for the PCL and suggest that only two factors should be extracted.

Comparability coefficients for two-factor solutions of increasing obliquity were also calculated. The value of the coefficients increased steadily with the increasing obliquity of the solutions. The most oblique Harris-Kaiser independent cluster solution resulted in the highest coefficients, suggesting that an oblique transformation is most appropriate for these data (Everett, 1983).

The independent two-factor solutions for Samples 1, 3, 4, and 7 were in close agreement; all of the congruence coefficients between corresponding factors equaled or surpassed .83, close to the criterion of .85 used by Harpur et al. (1988) to indicate the equivalence of two factors. There was less agreement among these samples for the three- or four-factor solutions. However, Sample 2 did not produce a solution congruent with any of the other four samples. For this reason, it was not included in the pooled-factor solution described below.

The primary-factor pattern matrices for Samples 1, 3, 4, and 7 were used to obtain the pooled solution, shown in Table 1, which represents a best fit to the four sample correlation matrices. Examination of both the pooled and sample solutions revealed a clear two-factor structure very similar to that described for the PCL. As was the case for the PCL, the item *promiscuous sexual behavior* failed to load above .4 on either factor. The loadings for two items differed for the PCL-R solution: *many short-term relationships* and *criminal versatility* failed to load above .4 on Factor 2, as they had for the PCL. If anything, the former loaded on Factor 1. All three items demonstrated inconsistent loadings across the four sample solutions.

Correlations between the two factors ranged from .39 to .54 ($M = .48$) in the separate-sample solutions. The congruence coefficients between corresponding factors for the common pattern and sample patterns transformed into congruence with the common pattern were uniformly high (range = .94-.98). The congruence coefficients between the sample patterns themselves ranged from .82 to .94 ($M = .89$).

On the basis of the results reported above, we judged that Factors 1 and 2 could be defined by eight and nine PCL-R items, respectively. Scores on Factor 1 were derived by summing the scores on Items 1, 2, 4, 5, 6, 7, 8, and 16. Scores on Factor 2 were derived by summing the scores on Items 3, 9, 10, 12, 13, 14, 15, 18, and 19.

The interrater reliability and internal consistency of the factors remained high despite the small number of items included in each scale. The mean ICCs for Factors 1 and 2 were .77 and .85, respectively, for a single rating, and .87 and .92, respectively, for the average of two ratings. The mean coefficient alphas were .84 and .79 for Factors 1 and 2, respectively. The corresponding mean interitem correlations were .39 and .31. Although the ICCs were slightly higher for Factor 2, which is scored more on the basis of objectively defined criteria than for Factor 1, which relies more on judgments and inferences about personality,

Factor 1 generally demonstrated higher internal consistency than Factor 2, as indicated by both alpha coefficients and mean interitem correlations.

Correlation Between PCL and PCL-R

The empirical factor structure and the nature of the scoring changes introduced suggest that the PCL-R measures the same construct that is measured by the PCL. However, more direct evidence may be obtained by comparing PCL total scores and factor-scale scores directly. So far, we have been able to obtain independent ratings on the PCL and PCL-R only with Sample 2. The correlation between the PCL and the PCL-R total scores was .88; correlations of .82 and .81 were obtained for Factors 1 and 2, respectively. However, these values are limited by the reliability of the scales themselves. True estimates of the correlations between the constructs measured by the PCL and the PCL-R are obtained when the correlations are disattenuated, or corrected for unreliability of the scales (Hakstian, Schroeder, & Rogers, 1988). Using ICC estimates of reliability, the true correlation between the PCL and the PCL-R total scores lies between .95 and 1.0.³ Correlations for Factor 1 lie between .94 and 1.0; those for Factor 2 lie between .93 and 1.0.

Discussion

The PCL is a reliable and valid measure of psychopathy in male criminal populations. The current results demonstrate the reliability and concurrent validity of the PCL-R. In addition, the results of a number of recent studies (described in the forthcoming manual) confirm the construct and predictive validity of the revised version. We have no hesitation in concluding that both versions measure the same construct and that findings obtained with one version are generalizable to the other. Both versions provide continuous scores reflecting the extent to which an individual matches the prototypical psychopath, as well as factor scores that permit the social deviance component of psychopathy to be separated from the cluster of personality traits that are fundamental to the construct.

In addition to validating the PCL-R, these results also provide a strong replication of an earlier study on the factorial structure of the PCL (Harpur et al., 1988). Of the 11 samples analyzed to date ($N = 2,119$), the optimal two-factor solution obtained by independent analysis of each sample has been highly consistent for 10 of the samples. We feel confident that the eight- and nine-item scales described here are adequate measures of the two factors and that results reported for the PCL factors will be applicable to the PCL-R factors. Interpretation of the factors as measuring a selfish, callous, and remorseless use of others (Factor 1) and a chronically unstable, antisocial, and socially deviant lifestyle (Factor 2) remains unchanged.

Sample 2 did not possess the same factor structure as the other samples in this study. There are several possible explana-

³ The precise value depends on whether the estimate of reliability for one rater or for two is used for disattenuation. Because some of the subjects in each sample were rated by two raters, these values represent upper and lower bounds for the reliabilities.

tions for this. The relatively small sample size ($n = 122$) may have resulted in a discrepant and unreliable factor structure resulting from sampling error. In addition, the data for this sample were gathered on provincial prison inmates, for whom far less extensive or complete file information was available. Finally, because this was the first sample with which we used the PCL-R, it is possible that minor changes in details of the scoring procedure were introduced as the data were gathered—changes that could have led to somewhat variable ratings for some items.

Despite the fact that the factor structure of Sample 2 was anomalous, the psychometric properties of the factor-scale scores derived from this sample were similar to those for the other samples. These scores were essentially identical to factor-scale scores independently derived from the PCL.

It is clear that the PCL, in both of its versions, is a robust measure of psychopathy in White, male criminals. Research is currently underway to assess the applicability of the PCL to Black, Native American, and female criminals, as well as to noncriminals.

References

- Côté, G., & Hodgins, S. (1989). *Une mesure de psychopathie validation de la version français* (The Psychopathy Checklist-French translation). Manuscript in preparation.
- Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research*, 18, 197–218.
- Hakstian, A. R., Schroeder, M. L., & Rogers, W. T. (1988). Inferential procedures for correlation coefficients corrected for attenuation. *Psychometrika*, 53, 27–43.
- 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. (1985). *The Psychopathy Checklist*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.
- Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the Psychopathy Checklist. *Journal of Consulting and Clinical Psychology*, 56, 741–747.
- Harpur, T. J., & Hare, R. D. (1989). *The assessment of psychopathy as a function of age*. Manuscript submitted for publication.
- Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 6–17.
- Hart, S. D., & Hare, R. D. (1989). Discriminant validity of the Psychopathy Checklist in a forensic psychiatric population. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 211–218.
- Meredith, W. (1964). Rotation to achieve factorial invariance. *Psychometrika*, 29, 187–206.
- Rice, M. E., & Harris, G. T. (1989). *A longterm study of psychopaths treated in a therapeutic community*. Manuscript submitted for publication.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Wong, S. (1988). Is Hare's Psychopathy Checklist reliable without an interview? *Psychological Reports*, 62, 931–934.

Received June 22, 1989

Revision received January 4, 1990

Accepted January 23, 1990 ■