Psychological Assessment
2001, Vol. 13, No. 2, 171–188

1040-3590/01/$5.00 DOI: 10.1037//1040-3590.13.2.171

Refining the Construct of Psychopathy: Towards a Hierarchical Model

David J. Cooke and Christine Michie
Glasgow Caledonian University

Psychopathy is characterized by diverse indicators. Clinical accounts have emphasized 3 distinct facets: interpersonal, affective, and behavioral. Research using the Psychopathy Checklist—Revised (PCL–R), however, has emphasized a 2-factor model. A review of the literature on the PCL–R and related measures of psychopathy, together with confirmatory factor analysis of PCL–R data from North American participants, indicates that the 2-factor model cannot be sustained. A 3-factor hierarchical model was developed in which a coherent superordinate factor, Psychopathy, is underpinned by 3 factors: Arrogant and Deceitful Interpersonal Style, Deficient Affective Experience, and Impulsive and Irresponsible Behavioral Style. The model was cross-validated on North American and Scottish PCL–R data, Psychopathy Screening Version data, and data derived from the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) antisocial personality disorder field trial.

Psychopathy is an important psychological construct. There is compelling evidence for its utility (see, e.g., Cooke, Forth, & Hare, 1998; Hare, 1991, 1996). Within applied settings, psychopathy is an important predictor of criminal behavior, particularly violence (Hart & Hare, 1997; Salekin, Rogers, & Sewell, 1996). Psychopathy is linked to failure on conditional release, violent recidivism, and poor treatment response (see, e.g., Hemphill, Hare, & Wong, 1998; Rice & Harris, 1992; Serin, 1996). Within the laboratory, psychopathy has been linked to a wide range of psychophysiological processes (see, e.g., Hare, 1978; Hare, Cooke, & Hart, 1999; Intrator et al., 1995; Lapierre, Braun, Hodgins, & Toupin, 1995; Newman & Wallace, 1993; Patrick, Bradley, & Lang, 1993).

The viability of a psychopathological construct is based on a range of evidence. A prerequisite is the existence of a coherent syndrome, that is, a cluster of symptoms, signs, and traits that occur together and that are distinct from other clusters (see, e.g., Blashfield & Draguns, 1976; Eysenck, 1970; E. Robins & Guze, 1970). The literature reveals that the defining characteristics of psychopathy are remarkably diverse; they entail interpersonal, behavioral, and affective traits (Cleckley, 1976; Hare, 1991; Lilienfeld, 1994; see Table 1 for a list of some defining characteristics). The very diversity of these traits may be contraindicative of a coherent syndrome.

Measuring Psychopathy With the Psychopathy Checklist—Revised

The Psychopathy Checklist—Revised (PCL–R) is the measure of choice within the field; however, there is a lack of clarity about the structural model that underpins PCL–R ratings. The PCL–R manual contains three putative structural models: a three-facet model, a two-factor model,[1] and a hierarchical[2] model (Hare, 1991). The dominant model in the PCL–R literature—and increasingly in related literatures—is the two-factor model (see, e.g., Cooke, 1995; Harpur, Hare, & Hakstian, 1989; Kosson, Smith, & Newman, 1990; Livesley, 1998; Salekin, Rogers, & Sewell, 1997). PCL–R items are considered to be underpinned by two distinct but correlated factors: The affective and interpersonal items coalesce to form "the selfish, callous, and remorseless use of other" factor (Hare, 1991, p. 76; see Table 1). The behavioral items coalesce to form "the chronically unstable and antisocial lifestyle; social deviance" factor (Hare, 1991, p. 76).

The Origins of the Two-Factor Model

The dominance of the two-factor model can be traced back to the influential work of Harpur, Hakstian, and Hare (1988). Fundamental to their analyses and many subsequent analyses (e.g., Cooke, 1995; Hobson & Shine, 1998; Kosson et al., 1990) is the

David J. Cooke and Christine Michie, Department of Psychology, Glasgow Caledonian University, Glasgow, Scotland.

We thank Robert Hare, Joseph Newman, Thomas Widiger, and Marnie Rice and her colleagues for providing us with their raw data. Joseph Newman's data were collected with support from National Institute of Mental Health Grant NH37711. The Scottish data were collected with support from a grant from the Legal and Criminological Research Branch of the Scottish Office and a grant from the Scottish Office Chief Scientist's Office Grant KPR/OPR/18/5. Analyses were carried out with support from the Scottish Office Chief Scientist's Office Grant K/OPR/15/9/F4 and Economic and Social Research Council Grant L133222704. We thank Edward Wozniak for his support throughout the collection of the Scottish data, Robert Hare for his continued guidance and support, and Ronald Blackburn, Stephen Hart, Vernon Quinsey, and Thomas Widiger for comments on an earlier version of this article.

Correspondence concerning this article should be addressed to David J. Cooke, who is now Director of Forensic Psychology Services, Douglas Inch Centre, 2 Woodside Terrace, Glasgow G3 7UY, United Kingdom. Electronic mail may be sent to djcooke@rgardens.u-net.com.

[1] Mathematically, two correlated factors are equivalent to two factors and a superordinate factor. The same is true for three correlated factors.

[2] The term *hierarchical* has three distinct meanings in this literature; we use the term to refer to nested dimensions analogous, for example, to the theoretical model underpinning intelligence where facets of intelligence relate to verbal and performance intelligence, which in turn are linked to the higher order construct of general intelligence (Blashfield & McElroy, 1995).

Table 1
Items in the Hare Psychopathy Checklist—Revised and Their Location in the Traditional Two-Factor Structure

Factor 1: Interpersonal/Affective	Factor 2: Social Deviance	Items that do not load on the two factors
1. Glibness/superficial charm	3. Need for stimulation/proneness to boredom	11. Promiscuous sexual behavior
2. Grandiose sense of self-worth	9. Parasitic lifestyle	17. Many short-term marital relationships
4. Pathological lying	10. Poor behavioral controls	20. Criminal versatility
5. Conning/manipulative	12. Early behavioral problems	
6. Lack of remorse or guilt	13. Lack of realistic, long-term goals	
7. Shallow affect	14. Impulsivity	
8. Callous/lack of empathy	15. Irresponsibility	
16. Failure to accept responsibility for own actions	18. Juvenile delinquency	
	19. Revocation of conditional release	

use of a measure of factor similarity, namely, the *congruence coefficient.* Although widely used, it has been subject to substantial criticism. Many authors have warned against the exclusive use of this index of similarity, arguing that a more robust strategy is to estimate a variety of indices of factorial similarity (see, e.g., Barrett, 1986; Floyd & Widaman, 1995; Gorsuch, 1983; Van de Vijver & Leung, 1997).

The interpretation of values of the congruence coefficient is also a problem: "Values higher than .95 are seen as evidence for factor similarity, whereas values lower than .90 (Van de Vijver & Poortinga, 1994) or .85 (Ten Berge, 1986) are taken to point to *nonnegligible incongruities* [italics added]" (Van de Vijver & Poortinga, 1994, p. 192).

Harpur et al. (1988) obtained data from six samples that ranged in size from 89 to 319 ($M = 187$). Using the procedure of split-half cross-validation developed by Everett (1983), they argued that a two-factor solution is most appropriate. They concluded that their analysis revealed "highly congruent factors between Samples 1–5" (Harpur et al., 1988, p. 743). Examination of the relevant table indicates that this conclusion is questionable. Nineteen (59%) of the 32 coefficients of congruence are below .90, the value that indicates "nonnegligible incongruities" (Van de Vijver & Poortinga, 1994, p. 192); over a third (34%) are below Ten Berge's (1986) less stringent cutoff of .85.

These analyses represent an improvement over previous approaches; bigger samples were used, and overfactoring was avoided. We consider that the paper by Harpur, Hakistan, and Hare (1988)—despite its profound influence on the field—does not provide compelling evidence for the adequacy of a two-factor model for psychopathy. Hare et al. (1990) used the same analytic approach and reached the same conclusions using PCL–R data.

The Establishment of the Two-Factor Model as a Gold Standard

Since the publication of Harpur et al.'s (1988) article, a series of studies has endeavored to replicate the analytic strategy, many using congruence coefficients to compare obtained solutions with the gold standard two-factor model. Kosson et al. (1990) collected Psychopathy Checklist (PCL) data on a sample of Caucasian ($N = 230$) and a sample of African American ($N = 123$) male prisoners. They assessed the congruence of the respective factors across samples, the value for Factor 1 being .67 and for Factor 2 being .93. These values suggest poor cross-sample generalization within this study. Other studies that have claimed to support the two-factor model include Cooke (1995; Scottish prisoners), Pham (1998; Belgian male high-security prisoners), Côté and Hodgins (1989; French-speaking Canadian prisoners), and Hobson and Shine (1998; English male prisoners). Recently, McDermott et al. (2000) replicated the two-factor model in a sample of prisoners but, using an analysis based on congruence coefficients and coefficients of comparability, argued in support of a unidimensional model for substance-dependent patients. These studies have used the PCL/PCL–R with adult male participants; several other studies have examined the factor structure in other populations.

Salekin et al. (1997) collected PCL–R data on 103 female offenders. They extracted two factors; seven of the expected eight variables loaded on Factor 1 but only four of the nine expected variables loaded on Factor 2.

Brandt, Kennedy, Patrick, and Curtin (1997) administered a modified version of the PCL–R suitable for adolescents to 130 males between 13 and 19 years of age. They obtained congruence coefficients of .91 and .84 for Factor 1 and Factor 2, respectively. They concluded that their results implied similarity between the structures found in adolescents and in adults.

Other workers have used measures other than PCL-derived measures and have argued that two factors similar to the gold standard factors can be discerned in their data. In a series of studies, Livesley (Livesley, Jackson, & Schroeder, 1992; Livesley & Schroeder, 1991) factor analyzed self-report data relating to antisocial personality disorder (ASPD) collected from general population, patient, and twin register samples. These studies yielded two factors that, the authors argued, are consonant with those obtained by Harpur et al. (1988).

The Application of Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) has been used in three studies. The use of CFA avoids two problems inherent in exploratory factor analysis (EFA): first, how many factors to extract and, second, how to rotate the factors extracted (Watson, Clark, & Harkness, 1994).

Hart, Cox, and Hare (1995) conducted CFA with the standardization sample of the Psychopathy Checklist: Screening Version (PCL:SV). Their results appear to support the two-factor model; however, no attempt was made to compare this model with other potential models. Brandt et al. (1997) conducted CFA on data derived from the modified version of the PCL–R suitable for

adolescents. They assessed the fit of the standard two-factor model. They obtained a comparative fit index (CFI) of .83 and reached the unwarranted conclusion that this indicated a "moderate fit with the predicted factor structure" (Brant et al., 1997, p. 432). Darke, Kaye, Finlay-Jones, and Hall (1998) collected PCL–R data on a sample of 376 participants in Australia. The two-factor simple structure did not fit their data.

Clinical Relevance of the Structure of Psychopathy

The two-factor model enhances assessment, management, and treatment of the disorder; key clinical variables have been shown to be differentially linked to the two factors. Hemphill et al.'s (1998) meta-analytic study demonstrated that Factor 2 is a stronger predictor of general recidivism than Factor 1, whereas both factors are equally important predictors of violent recidivism. Hart and Dempster (1997) found that Factor 1 is linked with planned predatory violence, whereas Factor 2 is linked with spontaneous and disinhibited violence. Treatment compliance and outcome are linked to the structure. Hobson, Shine, and Roberts (2000) reported that Factor 1—rather than Factor 2—is linked with disruptive behavior during therapeutic groups. Poor therapeutic change has been associated with high Factor 1 (see, e.g., Hughes, Hogue, Hollin, & Champion, 1997; Seto & Barbaree, 1999). Seto and Barbaree (1999) found that the highest rate of recidivism among sex offenders is observed among those who are both high in Factor 1 and rated by their therapist as having made good therapeutic progress. Understanding the structure of a disorder can assist both the selection of treatment targets and the management of treatment-interfering behaviors. Refining understanding of the structure of psychopathy should assist clinical intervention.

In summary, examination of both the origins of the two-factor model and subsequent attempts to replicate the model leaves us unconvinced of the adequacy of this model. The model appears to be founded on a misinterpretation of congruence coefficients. In the studies to be described, we reexamined the adequacy of the two-factor model of psychopathy. We developed a new hierarchical three-factor model of psychopathy and cross-validated it on North American PCL–R data, Scottish PCL–R data, and three other distinct measures of psychopathy or cognate constructs.

Study 1: Preliminary Factor Analysis

Introduction

In this study, we applied both EFA and CFA techniques to a large sample of North American PCL–R data to evaluate the validity of the two-factor model.

Method

Participants. Study 1 used PCL–R data from 2,067 North American participants derived from both correctional and forensic psychiatric settings. Approximately three quarters of the participants were part of the PCL–R standardization sample (Hare, 1991). Eight Canadian and two American samples were obtained. These samples were convenience samples collected for a range of clinical and experimental purposes. The Canadian samples included four samples of forensic patients (80 consecutive remands to a forensic hospital in British Columbia; 163 patients in the forensic unit of Pentanguishene Hospital Ontario; 132 patients in the Regional Psychiatric Center in Saskatoon, Saskatchewan; and 65 patients of a forensic outpatient clinic in Vancouver) and four prison samples (106 prisoners assessed in the Institute Phillipe Pinel de Montreal, 121 inmates at Oakalla provincial prison in British Columbia, 322 inmates at Matsqui federal medium security institution in British Columbia, and 87 inmates of a medium security prison in Kingston, Ontario). These samples were part of the standardization samples used in the PCL–R manual (Hare, 1991). In addition, the standardization sample included data from adult male prisoners in a minimum security institution in Wisconsin; these cases, as well as additional cases from the same source, were obtained for these analyses. The Wisconsin data consisted of a sample of 838 Caucasian prisoners and 153 African American prisoners. These data are described in detail elsewhere (Cooke & Michie, 1997, 1999).

Materials. The 20 PCL–R items, each reflecting a different trait or characteristic of psychopathy, are listed in Table 1. Information about the participant was obtained from interview and file review. Items are defined in detail in the PCL–R manual (Hare, 1991); the trained rater assesses the extent to which the characteristics described in the item definitions apply to the participant. Items are rated on a 3-point scale (0 = *item doesn't apply,* 1 = *item applies somewhat,* 2 = *item definitely applies*). The items are summed to yield a total score ranging from 0 to 40; this score reflects the degree to which an individual resembles the prototypical psychopath.

Results

We conducted an EFA of the standardization sample (N = 1,389; Hare, 1991) using EQS (Bentler & Wu, 1995). Comrey and Lee (1992) strongly advocated visual inspection of factor plots prior to rotation. When we inspected the loading plot for the two-factor solution, it was evident that no simple structure was present in these data. The points all lie more or less in an arc of 90°; thus, there is no way in which two axes—either orthogonal or oblique—could pass through, or close to, the majority of the points. Rotation was not carried out because this would have been misleading. Computerized rotational methods always force a solution even in the absence of true simple structure. If, as we have observed, many data sets appear to be described by similar models, this is merely because the correlation matrices are similar.

We then carried out CFA using EQS (Bentler & Wu, 1995) to determine whether the generally accepted two-factor model fitted the data. A simple factor structure was modeled in EQS with eight variables loading on Factor 1 and nine variables loading on Factor 2 as in the traditional model (Harpur et al., 1989). The quality of fit was estimated using multiple measures of fit because each measure has limitations and there are no agreed methods for absolutely determining goodness of fit (Kline, 1998).

The two-factor model did not provide an acceptable fit to the data, $\chi^2(118, N = 1{,}389) = 1{,}622.7, p < .001$, Akaike information criterion (AIC) = 1,386.7, Consistent AIC (CAIC) = 650.8, normed fit index (NFI) = .77, non-normed fit index (NNFI) = .75, CFI = .78, goodness-of-fit index (GFI) = .86, adjusted goodness of fit index (AGFI) = .82, root mean square error of approximation (RMSEA) = .10. It is also noteworthy that a one-factor model provided an even poorer fit (cf. McDermott et al., 2000), $\chi^2(169, N = 1{,}262) = 2{,}423.9, p < .001$, AIC = 2,085.8, CAIC = 1,048.1, NFI = .67, NNFI = .64, CFI = .68, GFI = .81, AGFI = .76, RMSEA = .10.

Discussion

We conclude that although the two-factor model has served as a useful heuristic device to guide research on psychopathy, it does not provide an adequate structural model for psychopathy.

Study 2: Developing a New Model

Introduction

In Study 2, we endeavored to develop a model that provides a more adequate fit to the data than the two-factor model. As suggested by Byrne (1994), we developed a new model taking into account both theoretical considerations regarding the nature of psychopathy and a combination of statistical techniques that explicate the dimensional structure underpinning PCL–R ratings.

Theoretical considerations in the model-building process. We took three broad theoretical consideration into account when developing the model.

1. The clinical tradition emphasizes three domains of defining characteristics. Hare (1991), following Cleckley (1976), argued that psychopathy is underpinned by affective, interpersonal, and behavioral components. Although there is debate regarding the relative importance of these three domains, there is agreement that the salient defining characteristics of the disorder are drawn from these three domains (American Psychiatric Association, 1994; Blackburn, 1998; Lilienfeld, 1994; World Health Organization, 1992).

2. Models of normal personality are hierarchical (see footnote 2). The dominant models in the field of normal personality are hierarchical in structure (see, e.g., Costa & McCrae, 1992; Eysenck, 1947; Watson et al., 1994). Hierarchical structures both describe and summarize the interrelations among items; items coalesce into traits with a higher order factor accounting for both the shared commonality of these traits and correlations among them (Watson et al., 1994; Zinbarg & Barlow, 1996).

3. Personality disorders may be best represented hierarchically because they are maladaptive forms of common personality traits. Widiger (Widiger & Lynam, 1998) has argued that psychopathy can be understood by reference to the five-factor model of normal personality. Floyd and Widaman (1995) argued that hierarchical approaches to measurement of clinically important constructs are underused. These approaches are appropriate because "most psychological constructs are composed of multiple, correlated facets" (Floyd & Widaman, 1995, p. 293).

Statistical considerations in the model-building process. In evaluating the structural properties of the disorder, we applied both a top-down and a bottom-up approach. Following Comrey and Lee (1992), we conducted EFA using EQS. The initial factors were rotated using the direct oblimin criterion to obtain an oblique rotated solution (Bentler & Wu, 1995). Five factors with eigenvalues greater than one were rotated. Comrey and Lee argued that this procedure produces the maximum number of legitimate factors possible and serves as a good starting point for EFA. Specific factors can be combined or excluded in later analyses.

Method

Participants. The complete sample available in Study 1 (N = 2,067) was randomly divided into two sets using SPSS (SPSS, 1993) so that any model developed on the first data set could be cross-validated on the second data set. Cross-validation is an important procedure for assessing the validity of a structure in structural equation modeling (Van de Vijver & Leung, 1997).

Materials. The PCL–R ratings described in Study 1 were used.

Results

Examination of Table 2 indicates that many of the items in the original two-factor structure load on two factors in these analyses, namely, Factor 1 and Factor 4. The remaining three factors were defined by pairs of items, the similarity of item content producing highly specific factors. Factor 2 was defined by early behavioral problems and juvenile delinquency, Factor 3 was defined by pro-

Table 2
Factor Loadings of Five Factors With Eigenvalues Greater Than One Following Oblique Rotation

Item	Factor 1	2	3	4	5
1. Glibness/superficial charm	.39	−.20	.27	−.10	.35
2. Grandiose sense of self-worth	.40	−.21	.23	−.02	.31
3. Need for stimulation/proneness to boredom	.05	.10	.08	.44	.21
4. Pathological lying	.32	−.09	.02	.13	.36
5. Conning/manipulative	.21	−.04	.20	.15	.32
6. Lack of remorse or guilt	.66	.08	.03	.08	−.08
7. Shallow affect	.57	.15	.03	.10	.01
8. Callous/lack of empathy	.59	.18	.14	.14	−.10
9. Parasitic lifestyle	−.05	−.02	−.03	.67	−.06
10. Poor behavioral controls	.32	.39	.08	.07	−.04
11. Promiscuous sexual behavior	.04	.08	.59	−.02	−.01
12. Early behavioral problems	.10	.53	.07	.10	.11
13. Lack of realistic, long-term goals	.11	.03	−.01	.53	.00
14. Impulsivity	.04	.13	.02	.52	.05
15. Irresponsibility	.07	−.10	.14	.55	.03
16. Failure to accept responsibility for own actions	.59	−.11	−.16	.10	.07
17. Many short-term marital relationships	−.12	−.03	.64	.09	−.02
18. Juvenile delinquency	−.08	.51	.01	.07	.31
19. Revocation of conditional release	−.12	.10	−.01	.13	.56
20. Criminal versatility	.05	.21	.09	−.04	.48

miscuous sexual behavior and many short-term marital relationships, and Factor 5 was defined by revocation of conditional release and criminal versatility. Items 1, 4, 5, and 10 failed to load significantly on any factor. This is essentially a top-down approach to the description of the dimensional structure; we then engaged in further exploration of the structure using a bottom-up approach.

We examined the structure of the PCL–R ratings using a technique derived from item response theory (IRT; Embretson, 1996). In IRT, local dependence occurs when two or more items are more highly associated than can be explained by their relationships with the underlying latent trait (Steinberg & Thissen, 1996). Local dependence can occur because the same information is used to score more than one item, and thus, a pair of items may actually represent somewhere between one and two items.[3] Failure to take account of local dependence can lead to a lack of clarity in the structural model. There are good grounds for assuming that local dependence may be a problem with the PCL–R. A straightforward example of content overlap is where an incident of escape can contribute to scores on two items, namely, revocation of conditional release and criminal versatility. Examination of extracts from the rating descriptions for PCL–R items reveals a range of more subtle overlaps in content.

There are no universally agreed-on criteria for local dependence. However, Q_3 has been shown to be a robust measure (Chen & Thissen, 1997).[4] Positive values of .15 or greater are suggestive of some local dependency. All items other than poor behavioral controls displayed local dependency with at least one other item; they formed what have been termed testlets (Steinberg & Thissen, 1996). For example, item pairs including shallow affect and callous/lack of empathy (Q_3 = .22) and early behavioral problems and juvenile delinquency (Q_3 = .22) formed testlets. Three items, need for stimulation/proneness to boredom, impulsivity, and irresponsibility, formed a three-variable testlet.

To examine these relationships more systematically, we used these Q_3 values as proximity measures in a group-average agglomerate cluster analysis (Everitt, 1993). Figure 1 is a dendogram in which the higher the level at which items fuse, the greater is the similarity of items. This figure indicates that the items agglomerate into three distinct clusters of items. Of particular note is the observation that traditional Factor 1 items form two discrete clusters of items.

Building the model. We then developed a new model iteratively by taking into account both the theoretical and the statistical considerations alluded to above. Our starting point was the model that fitted five factors with eigenvalues greater than one; 17 of the 20 items had loadings of .40 or greater, and three of the factors were specific factors produced by pairs of items that formed individual testlets. The model was fitted using EQS. Examination of Table 3 indicates that the model provides an adequate fit.

In the next step of the model-building process, we estimated a model containing the two major dimensions of variation identified above, namely, Factors 1 and 4. Consideration of the fit indices in Table 3 indicates that this 10-item model provides an acceptable fit.

On the basis of both theoretical considerations and the empirical findings from the testlet analyses described above, we then modified the above model by adding another level to the hierarchy, namely, a level that specified the testlets. Consideration of the fit indices in Table 3 indicates that this model provides a good fit. In particular, the AIC and CAIC indices imply that it represents a substantial improvement over the prior 10-item model in terms of parsimony; this improvement is statistically significant, $\Delta\chi^2(6, N = 1{,}022) = 89.7, p < .001$.

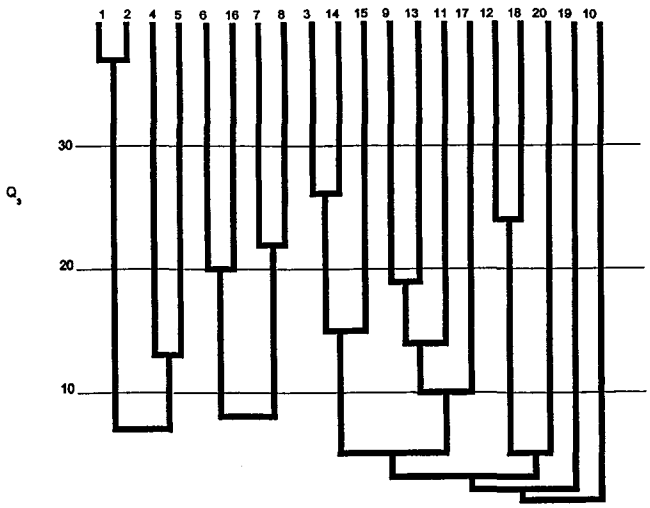

Figure 1. Group-average agglomerate cluster analysis of Q_3 values for Psychopathy Checklist—Revised items.

Unfortunately, this 10-item model does not include 3 items thought to be of central diagnostic significance to the disorder, namely, glibness/superficial charm, pathological lying, and conning/manipulative. Both theoretical considerations about the nature of psychopathy and statistical evidence of the importance of these items (Cooke & Michie, 1997, 1999) suggest that there is an imperative to include them in the model. From a statistical perspective, IRT analysis has demonstrated that these 3 items provide substantial amounts of information compared with the 7 items that remain excluded from the model. We estimated a 13-item model in which these 3 items were added within the testlets that underpinned the first factor. Consideration of the fit indices in Table 3 indicates that this model provides an adequate fit, although not as good a one as the 10-item model.

In the final step in the model-building process, we took into account the findings from the cluster analysis of the items. This analysis indicated that the items that traditionally have been represented as Factor 1 items contain two distinct clusters of items. We introduced a third factor into the model. The resultant model thus contains 13 items and six testlets with three factors loading onto a higher order factor. This model provides a good fit (Table 3). In particular, the drop in the AIC and CAIC indices together with $\Delta\chi^2$ imply that this model represents a substantial improvement over the prior two-factor hierarchical model containing 13 items, $\Delta\chi^2(1, N = 1{,}018) = 131.8, p < .001$.

[3] In principle, local dependence could be assessed using factor analytic approaches; however, the number of factors that have to be fitted is large, and this generally results in unstable solutions (Steinberg & Thissen, 1996).

[4] Q_3 is the correlation of the residuals from the IRT model, $d_{ik} = u_{ik} - \hat{P}_i(\hat{\theta}_k)$ and $Q_{3ij} = r_{d_i d_j}$, where u_{ik} is the score of the kth test taker on the ith item. A table of Q_3 values is available from the authors.

Table 3
Goodness-of-Fit Indices for the Building of the Model

Model	N	χ^2	df	AIC	CAIC	NFI	NNFI	CFI	GFI	AGFI	RMSEA
5 factors from EFA, cutoff 0.4	799	351.8	94	163.8	−370.4	.90	.90	.92	.95	.92	.06
2 correlated factors (1 and 4 from EFA)	1,022	202.3	34	134.3	−67.3	.93	.92	.94	.96	.94	.07
10 items, testlets, 2 correlated factors	1,022	112.6	28	56.6	−109.4	.96	.95	.97	.98	.96	.05
13 items, testlets, 2 correlated factors	1,018	357.9	57	243.9	−93.8	.92	.90	.93	.95	.92	.07
13 items, testlets, 3 correlated factors	1,018	226.1	56	114.1	−217.8	.95	.94	.96	.97	.95	.05

Note. AGFI = Adjusted Goodness of Fit Index; AIC = Akaike Information Criterion; CAIC = Consistent Akaike Information Criterion; CFI = comparative fit index; EFA = exploratory factor analysis; GFI = goodness-of-fit index; NFI = normed fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation.

The final model with completely standardized parameter estimates is displayed in Figure 2. All PCL–R items had significant loadings on the testlets ($ps < .001$), ranging from .58 to .83 (arithmetic mean loading = .70). The testlets all had significant loadings on the first-order factors ($ps < .01$), ranging from 0.81 to 1.00 (arithmetic mean loading = .90). Finally, the three factors loaded significantly on the superordinate factor ($ps < .01$), ranging from .82 to .87 (arithmetic mean loading = .84).

Examination of Factor 1 indicated that it measures interpersonal style being specified by two testlets, the first being defined by the items glibness and superficial charm and grandiose sense of self-worth and the second being defined by the items pathological lying and conning/manipulative. Factor 1 may be described as Arrogant and Deceitful Interpersonal Style. Factor 2 represents an affective factor, being specified by two testlets, the first defined by the items shallow affect and callous/lack of empathy and the second defined by lack of remorse or guilt and failure to accept responsibility. Factor 2 may be described as Deficient Affective Experience. Factor 3 represents a behavioral factor specified by two testlets, the first defined by three items (need for stimulation/proneness to boredom, impulsivity, and irresponsibility) and the second defined by parasitic lifestyle and lack of realistic, long-term goals. Factor 3 may be described as Impulsive and Irresponsible Behavioral Style. Given that all these first-order factors contribute to a higher order factor, we consider that this higher order factor can be defined as Psychopathy.

Refining the model. Having established a basic model, we tested it to determine first, whether it could be modified to provide a more parsimonious model, and second, to determine whether the interposing of the traditional Factor 1, namely, "the selfish, callous, and remorseless use of other" (Hare, 1991, p. 76), as an intermediate step between two new factors, namely, Arrogant and Deceitful Interpersonal Style and Deficient Affective Experience, and the superordinate Psychopathy factor would improve the quality of fit. Third, we examined the extent to which further restraints could be imposed on the model by testing whether loadings could be made equal. The competing models were nested, and comparative fit was tested using nested chi-square tests (Byrne, 1994). The fit indices for the models tested are displayed in Table 4.

We estimated three models to determine whether different levels in the hierarchical model could be deleted—one at a time—to improve parsimony without a degradation of overall fit. First, we removed the testlet level; examination of the fit indices indicated that this model provided an adequate fit. However, a nested chi-square test indicated that the fit was significantly poorer than that achieved with the original model, $\Delta\chi^2(6, N = 1{,}018) = 165.1, p < .001$. Second, we removed the factor level; examination of the fit indices revealed that it provided an adequate fit to the data. However, a nested chi-square test indicated that the fit was significantly poorer than that achieved with the original model, $\Delta\chi^2(3, N = 1{,}018) = 198.6, p < .001$. Third, we removed the superordinate Psychopathy factor (i.e., the three first-order factors were set to be uncorrelated). The fit indices indicated that this model did not provide a good fit to the data, and a nested chi-square test indicated that this model degraded fit significantly, $\Delta\chi^2(3, N = 1{,}018) = 775.2, p < .001$. These analyses, taken together, suggest that all levels of the hierarchy are necessary when describing this disorder.

Because of the theoretical predominance of the two-factor model, we attempted to fit the traditional Factor 1 as an intermediate step in the model; a nested chi-square test indicated that this model did not change the fit significantly, $\Delta\chi^2(1, N = 1{,}018) = .0$, $p = ns$. However, as can be observed from Table 4, a reduction in parsimony was evident from the AIC and CAIC indices. We concluded that the three-factor hierarchical model provides the best description of these data.

We then endeavored to restrict the model further by testing whether it was possible to constrain the loadings of the items on the testlets, the testlets on the factors, and the factors on the superordinate factor to be equal, respectively. The loadings of the items on the testlets ranged from .58 to .83. Constraining these loadings to be equal within testlets resulted in a significant degradation of the model, confirming significant variation across the loadings' fit, $\Delta\chi^2(7, N = 1{,}018) = 44.2, p < .001$. The loadings of the testlets on the factors ranged from .81 to 1.0. Constraining these loadings to be equal within factors also produced a significant degradation in fit, confirming that there was significant variation across these loadings, $\Delta\chi^2(3, N = 1{,}018) = 13.7, p < .01$. The loadings of the factors on the superordinate factor ranged from .82 to .87. Constraining these loadings did not result in a significant degradation in fit, $\Delta\chi^2(2, N = 1{,}018) = 4.3$, *ns*. This finding indicates that each of the factors contributes to the superordinate factor to a similar extent.

The final model, with three factors loading on a superordinate factor, is equivalent to a model in which the three factors are assumed to be correlated. The data were fitted to a model in which the factors were correlated; the factors were all highly intercorrelated, F_1vF_2, $r = .71$; F_1vF_3, $r = .68$; F_2vF_3, $r = .73$. All correlations were significantly different from 0 and 1 ($ps < .001$).

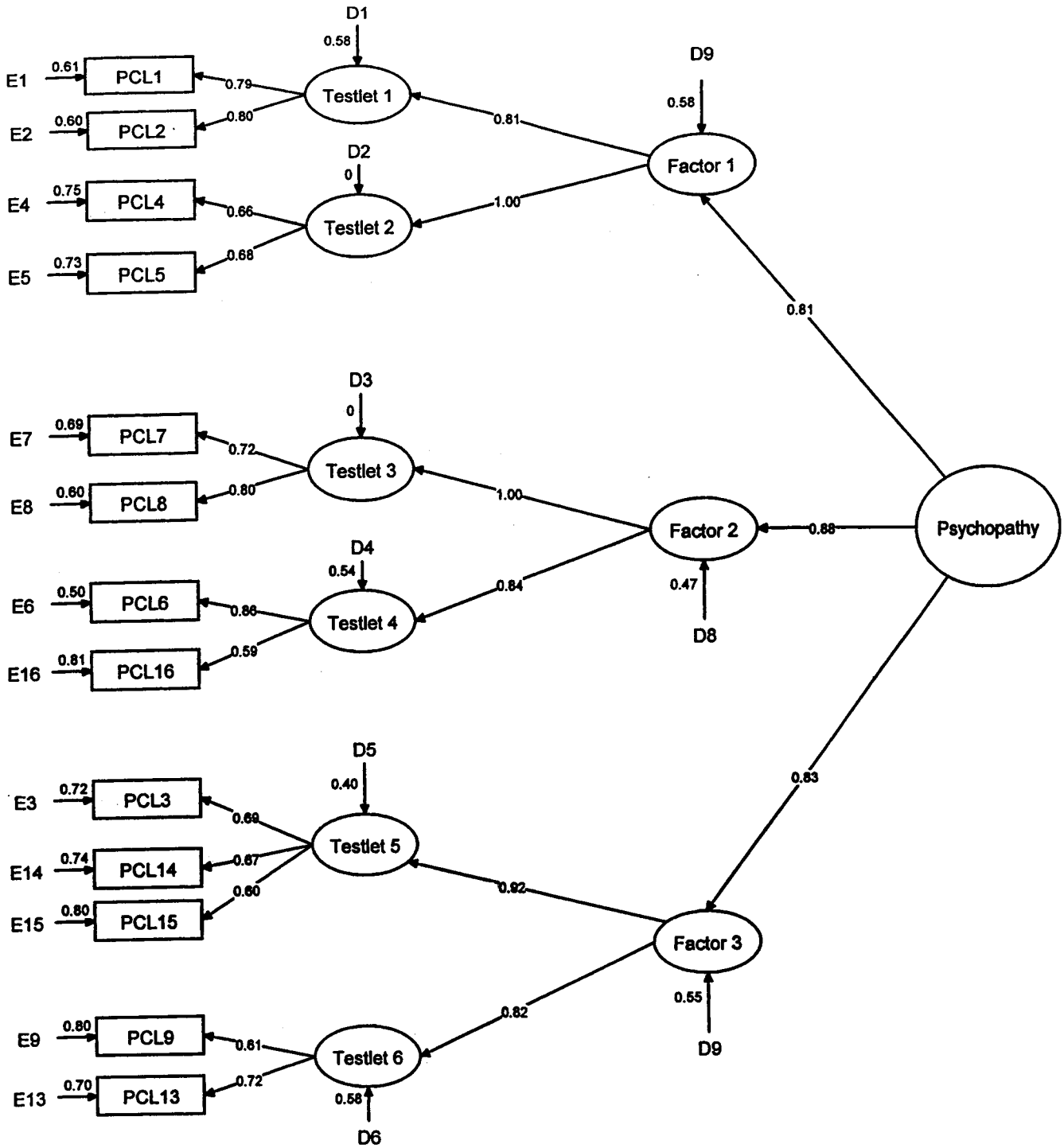

Figure 2. Standardized estimates for the hierarchical factor model derived by confirmatory factor analysis of Psychopathy Checklist—Revised items.

Thus, all three factors are important and are different from each other.

Having refined the model, we then proceeded to determine whether the superordinate factor could be regarded as unidimensional or coherent (Zinbarg & McDonald, 1999). We calculated the indirect effects of the individual items on the superordinate factor by multiplying the loadings on the paths between the item and the superordinate factor. This procedure provides an estimate of the total test variance accounted for by the superordinate factor: The ratio of this quantity to the observed variance in total scores

Table 4
Goodness-of-Fit Indices for Refinement of the Model

Model	*N*	χ^2	*df*	AIC	CAIC	NFI	NNFI	CFI	GFI	AGFI	RMSEA
Testlets, factors, superordinate factor	1,018	226.1	56	114.1	−217.8	.95	.94	.96	.97	.95	.05
Remove testlets	1,018	391.2	62	267.2	−100.2	.91	.90	.92	.94	.92	.07
Remove factors	1,018	424.7	59	306.7	−42.9	.90	.89	.92	.94	.91	.08
Remove superordinate factor (i.e., 3 uncorrelated factors)	1,018	1,001.3	59	883.3	533.7	.77	.71	.78	.87	.78	.12
Add level: combine Factors 1 and 2 to reproduce "Factor 1"	1,018	226.1	55	116.1	−209.8	.95	.94	.96	.97	.95	.05
Testlet loadings equal	1,018	270.3	63	144.3	−229.0	.94	.94	.95	.96	.94	.06
Factor loadings equal	1,018	239.8	59	121.8	−227.8	.95	.94	.96	.96	.95	.05
Superordinate factor loadings equal	1,018	230.4	58	114.4	−229.3	.95	.95	.96	.97	.95	.05

Note. AGFI = Adjusted Goodness of Fit Index; AIC = Akaike Information Criterion; CAIC = Consistent Akaike Information Criterion; CFI = comparative fit index; GFI = goodness-of-fit index; NFI = normed fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation.

provides the estimate of general factor saturation (GFS; Zinbarg & McDonald, 1999). The estimate of GFS indicated that the superordinate factor in this model is essentially a coherent construct explaining more than three quarters of the variance accounted for by the model (*GFS* = .77). Zinbarg, Barlow, and Brown (1997) indicated that values over .50 are consistent with a coherent construct.

Discussion

This study indicates that PCL–R data fit a three-factor hierarchical model; the superordinate construct is a coherent construct. The fit achieved is superior to the fit achieved with the two-factor model.

The plausibility and verisimilitude of a structural model are substantially enhanced by the cross-validation of the model using samples or measures that are independent of those on which the original model was based (Byrne, 1994; Van de Vijver & Leung, 1997). Failure to cross-validate means that misspecified models can emerge because of capitalization on chance factors or peculiarities of the sample on which the original model was based. The need to cross-validate, although fundamental, is more honored in the breach than in the observance (see, e.g., MacCallum, Roznowski, & Necowitz, 1992). In the following studies, we attempted to cross-validate the hierarchical model within the same population, within the Scottish population, across two cognate measures of the same construct, and using other diagnostic criteria.

Study 3: Cross-Validation Within the North American Sample

Introduction

In Study 3, we determined whether the model developed on a random sample of North American PCL–R ratings could be generalized to a second random sample of these data.

Method

As reported above, in Study 2, the North American data set was divided into two randomly. To test the replicability of the final model, we fitted it using the second half of the data set.

Results

Examination of the fit statistics revealed that a very satisfactory fit was achieved, $\chi^2(56, N = 976) = 178.2, p < .001$, AIC = 66.2, CAIC = −263.2, NFI = .96, NNFI = .96, CFI = .97, GFI = .97, AGFI = .96, RMSEA = .05. The values achieved on the standard fit indices are all marginally higher than those obtained for the original data on which the model was developed, indicating good cross-validation of the model.

We achieved a more powerful test of cross-sample invariance by fitting the 13-item hierarchical model to the data from the two random samples simultaneously. This analytic strategy allows determination of whether the factorial structure of the instrument replicates across independent samples. A baseline model in which all the parameters were allowed to take different values for the two samples was estimated. For this unconstrained baseline model, $\chi^2(112, N = 1{,}994) = 404.3, p < .001$. Estimating a model in which all model parameters (loadings and variances) were constrained to be equal for the two samples resulted in a nonsignificant change in chi-square, $\Delta\chi^2(33, N = 1{,}994) = 22.3$, *ns*. This indicates that the model fits identically in the two random samples.

Discussion

These results show that the good level of fit achieved in the first random sample cannot be attributed to capitalization on chance or peculiarities of that sample. This increases the plausibility of the hierarchical three-factor model.

Study 4: Cross-Validation From North American Data to Scottish Data

Introduction

Cross-validation of a structural model across cultures is a strong test of the validity of a model; however, examples in the literature are rare (Van de Vijver & Leung, 1997). Previous research using EFA and IRT methods has demonstrated a degree of cross-cultural generalizability for the PCL–R from North American samples to Scottish samples (Cooke, 1995; Cooke & Michie, 1999), the primary difference being that the interpersonal features of the

disorder are less evident in Scottish participants except in extreme cases of the disorder.

Method

Participants. Data from three Scottish prison studies were used. Sample 1 contained 247 male prisoners from a systematic random sample of prisoners in the Scottish prison system (this sample is described in greater detail elsewhere; e.g., Cooke, 1995; Cooke & Michie, 1999). Sample 2 contained 105 adult male prisoners who participated in a study of early childhood experiences of criminal psychopaths (Marshall & Cooke, 1999). Sample 3 contained a random sample of 244 adult male prisoners who participated in a study of violence in relation to psychopathy and other risk factors (Cooke, Michie, Philip, & Carr, 1997; Michie & Cooke, 2000). All participants were volunteers. Over 99% were Caucasian. The mean age was 27.9 (SD = 8.7).

Materials. In all three studies, the PCL–R was completed by trained raters.

Results

The model developed and cross-validated using the North American data was fitted to the Scottish data. Examination of the fit statistics indicated that the model provided a just adequate fit to the data; the value of CFI—Byrne's (1994) index of choice—was over .90, $\chi^2(56, N = 530) = 228.0$, $p < .001$, AIC = 116.0, CAIC = −179.3, NFI = .89, NNFI = .88, CFI = .91, GFI = .93, AGFI = .88, RMSEA = .08. The estimated GFS indicates that the superordinate factor in this model is essentially a coherent construct (GFS = .79).

To determine whether specific aspects of the model fit more adequately than others, we again adopted the strategy of estimating the 13-item hierarchical model using the data from North America and Scotland simultaneously; all complete cases for each data set were used. A baseline model was estimated in which all the parameters were allowed to take different values for the two samples. For this unconstrained baseline model, $\chi^2(112, N = 2{,}524) = 585.3$, $p < .001$. Constraining all loadings (i.e., item–testlet, testlet–factor, and factor–superordinate factor loadings) in the model to be equal for the two samples resulted in a significant change, $\Delta\chi^2(12, N = 2{,}524) = 76.2$, $p < .001$. This indicates that the model does not fit identically in the two samples. Empirical and statistical criteria were used to alter the constraints imposed on the model. Previous research (Cooke & Michie, 1999) indicated that items that load on the traditional first factor, "the selfish, callous, and remorseless use of other" factor (Hare, 1991, p. 76), become positive only at high levels of psychopathy in the Scottish sample as compared with the North American sample. Examination of the Lagrange multiplier test values (Byrne, 1994) confirmed that the constraint that contributed most to $\Delta\chi^2$ was that which constrained the loading of the new first factor on the superordinate factor to be equal for both samples, $\chi^2(1, N = 2{,}524) = 27.0$, $p < .001$. Other constraints that also contributed to $\Delta\chi^2$ were those relating to the item grandiose sense of self-worth loading on the testlet that it defines together with glibness/superficial charm, $\chi^2(1, N = 2{,}524) = 7.4$, $p = .006$, and, to a lesser extent, the item impulsivity loading on the testlet that it defines together with the items need for stimulation/proneness to boredom and irresponsibility, $\chi^2(1, N = 2{,}524) = 4.3$, $p = .039$. On the basis of these results and previous findings, we tested the hypothesis that all loadings within the Factor 2 and Factor 3 branches of the model could be constrained to be equal, and the loadings within the Factor 1 branch were allowed to vary. When we imposed this set of constraints, the resulting change in chi-square was nonsignificant, $\Delta\chi^2(8, N = 2{,}524) = 14.9$, *ns*. Previous analyses indicated that these differences are likely to be cultural differences rather than rater effects (Cooke, 1995; Cooke & Michie, 1999).

Discussion

The adequate fit achieved by the model when applied to these data indicates a reasonable degree of cross-cultural generalizability of the model. In line with previous research that used different methods and smaller samples, it appears that the differences relate to the interpersonal aspects of the disorder.

Study 5: Cross-Validation From the Psychopathy Checklist—Revised to the Psychopathy Checklist: Screening Version

Introduction

In Study 5, we endeavored to extend the cross-validation of the three-factor model by fitting a comparable model to a different instrument, namely, the PCL:SV.

Method

The Screening Version of the PCL–R was developed first, to reduce the time and effort required to make assessments as compared with the PCL–R, and second, to allow for ratings to be performed in settings where criminal records are unavailable and/or irrelevant (thus rendering several PCL–R items impossible to score; Cooke, Michie, Hart, & Hare, 1999; Hart et al., 1995).

Participants. The standardization samples are described in full in the PCL:SV manual. In total, there were 586 participants from 11 different samples collected in Canada and the United States. Participants came from one of four settings: forensic nonpsychiatric (i.e., convicted prisoners not identified as mentally ill), N = 149; forensic psychiatric (i.e., individuals charged with or convicted of offenses who were identified formally as mentally disordered and who were being treated either as in- or outpatients), N = 120; civil psychiatric (i.e., individuals formally identified as mentally disordered but with no current charges or convictions and being treated as in- and outpatients), N = 217; and civil nonpsychiatric (i.e., community residents—university students—not currently identified as mentally disordered and with no current charges or convictions), N = 100. The average age was 30.9 (SD = 8.6); 31% of the participants were female; 76% of the participants were Caucasian. The participants were volunteers (see Cooke et al., 1999; Hart et al., 1995; for further details).

Materials. The PCL:SV is a 12-item rating scale based directly on the PCL–R. The item descriptions in the PCL:SV manual are very brief compared with the descriptions for the PCL–R items; they require less detailed information to be scored. The majority of the PCL:SV items were developed from single PCL–R items by shortening and simplifying the items without losing their essential meaning. The remaining PCL:SV items were derived by first collapsing, then shortening and simplifying, pairs of PCL–R items that are highly similar in content (in essence, testlets). For example, PCL:SV Item 5 (lacks empathy) reflects a combination of PCL–R

Items 8 (callous/lack of empathy) and 7 (shallow affect).[5] It is the case that the developers, through experience, identified many of the testlets observed in PCL–R data and deleted many of the items that contribute little information to the estimate of the trait (Cooke et al., 1999).

Results

A model was developed with cognate PCL:SV items being substituted for PCL–R items in the original model. Our aim was to specify the same testlets in this model as had been specified in the PCL–R model. This model is displayed in Figure 3. All items that loaded on testlets were significant ($ps < .001$), all loadings on the factors were significant ($ps < .001$), and finally, all loadings of the factors on the superordinate factor were significant ($ps < .001$).

Each factor in the PCL:SV model was underpinned by one testlet and a singleton item. The Arrogant and Deceitful Interpersonal Style factor was underpinned by a testlet that combines the items superficial and grandiose and the singleton item deceitful. The Deficient Affective Experience factor was underpinned by a testlet that combines the items lacks remorse and doesn't accept responsibility and the singleton item lacks empathy. The Impulsive and Irresponsible Behavioral Style factor was underpinned by a testlet composed of the items impulsive and irresponsible and the singleton item lacks goals. These three factors loaded strongly on the superordinate factor (arithmetic mean loading = .91).

This model provides an excellent fit to these data, indicating that the structural model developed with PCL–R data can be generalized to the PCL:SV, $\chi^2(21, N = 586) = 41.6$, $p < .001$, AIC = −0.4, CAIC = −113.2, NFI = .98, NNFI = .99, CFI = .99, GFI = .98, AGFI = .97, RMSEA = .04. The estimated GFS indicates that the superordinate factor in this model is essentially a coherent construct (GFS = .85).

Discussion

The results of this study demonstrate that the three-factor hierarchical model generalizes to a related, yet distinct, method of measuring psychopathy, a method that uses fewer items and that requires less collateral information.

Study 6: Cross-Validation From the Psychopathy Checklist—Revised to the Psychopathy Criterion Set

Introduction

In Study 6, we endeavored to extend the cross-validation of the three-factor hierarchical model by fitting a comparable model to a different set of criteria, namely, the Psychopathy Criterion Set (PCS).

Method

Twelve field trials were carried out in the development of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–IV;* American Psychiatric Association, 1994); ASPD was the subject of one field trial that aimed to evaluate proposed changes to the definitions of the disorder (Widiger et al., 1996). A detailed account of the field trial is provided by Widiger et al. (1996).

Participants. Data from 506 participants were obtained in five settings in which the base rate of ASPD was likely to be high; these settings included inpatient, outpatient, and prison settings. These included 100 male inmates of a medium security penal institution in Canada. All other samples were drawn in the United States. These included 100 male and female inpatients of a general psychiatric hospital, 100 male and female outpatients receiving methadone maintenance, 101 people within drug treatment facilities or homeless shelters, and 106 adopted away offspring at risk from ASPD as a consequence of the disorder of their biological parents. Detailed demographic characteristics are provided in Widiger et al. (1996). The average age was 32.7 ($SD = 8.2$); 33% of the participants were female; 68% of the participants were Caucasian.

Materials. Each participant was interviewed using a semistructured interview designed to measure symptoms and signs from three diagnostic criteria sets. The first criteria set was the *Diagnostic and Statistical Manual of Mental Disorders* third edition, revised, ASPD symptoms (American Psychiatric Association, 1987). The second criteria set was derived from Hare's PCL–R (Hare, 1991). The PCS criteria set contains 10 characteristics derived from the PCL:SV. The third criteria set was the ICD–10 criteria set of seven characteristics used to measure dyssocial personality disorder (DPD; World Health Organization, 1992).

Results

Data from the ASPD field trial were made available by T. A. Widiger (personal communication, August 14, 1997). A model was developed with cognate PCS items being substituted for PCL–R items to form a model containing seven items, three singleton items, and two testlets, as illustrated in Figure 4. All loadings were significant ($ps < .001$).

The Arrogant and Deceitful Interpersonal Style factor was underpinned by a testlet that combined the items glib and superficial with inflated and arrogant self-appraisal and the singleton item deceitful and manipulative. The Deficient Affective Experience factor was underpinned by two singleton items, lacks remorse and lacks empathy. The Impulsive and Irresponsible Behavioral Style factor was underpinned by two singleton items, impulsive and irresponsible. These three factors loaded strongly on the superordinate factor (arithmetic mean loading = .92).

This model provides an excellent fit to these data, indicating that the structural model developed with PCL–R data can be generalized to the PCS, $\chi^2(10, N = 501) = 40.2$, $p < .001$, AIC = 20.2, CAIC = −32.0, NFI = .98, NNFI = .96, CFI = .98, GFI = .98, AGFI = .94, RMSEA = .08. The estimated GFS indicated that the superordinate factor in this model was essentially a coherent construct (GFS = .84).

Discussion

This further test of cross-validation provides additional confirmation of the generalizability of the three-factor hierarchical model.

---

[5] PCL:SV items with corresponding PCL–R item numbers in brackets: (1) superficial [1]; (2) grandiose [2]; (3) deceitful [4, 5]; (4) lacks remorse [6]; (5) lacks empathy [7, 8]; (6) doesn't accept responsibility [16]; (7) impulsive [3, 14]; (8) poor behavioral controls [10]; (9) lacks goals [9, 13]; (10) irresponsible [15]; (11) adolescent antisocial behavior [12, 18]; (12) adult antisocial behavior [19, 20]. Two PCL–R items, namely, many short-term marital relationships and promiscuous sexual behavior, were not used in the development of the PCL:SV.

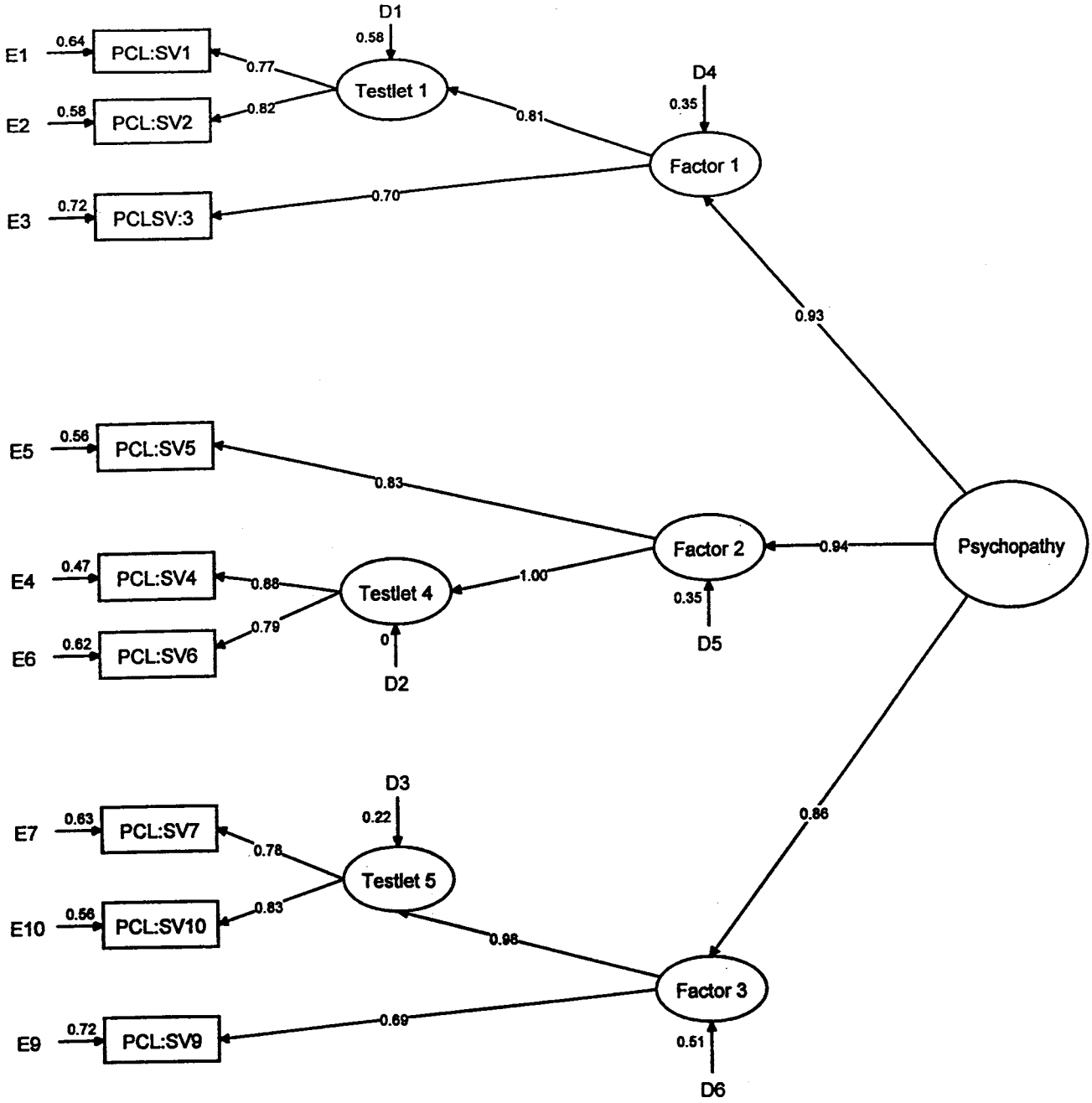

Figure 3. Standardized estimates for the hierarchical factor model derived by confirmatory factor analysis of Psychopathy Checklist: Screening Version (PCL:SV) items.

Study 7: Cross-Validation From the Psychopathy Checklist—Revised to Psychopathy Criterion Set, Antisocial Personality Disorder, and Dissocial Personality Disorder Criteria

Introduction

All the models derived so far have been based on PCL–R, or PCL–R derived, measures of psychopathy. To assess whether it was possible to generalize the model beyond these measures, we endeavored to develop a model using the adult ASPD and DPD criteria set from the *DSM–IV* ASPD field trial.

Method

Participants. The participants were the same as those in the previous study.

Materials. Initially, we had hoped to develop a model using only ASPD or DPD criteria, the PCS criteria being excluded. We fitted a model (see Figure 5) using the cognate variables and obtained some evidence of

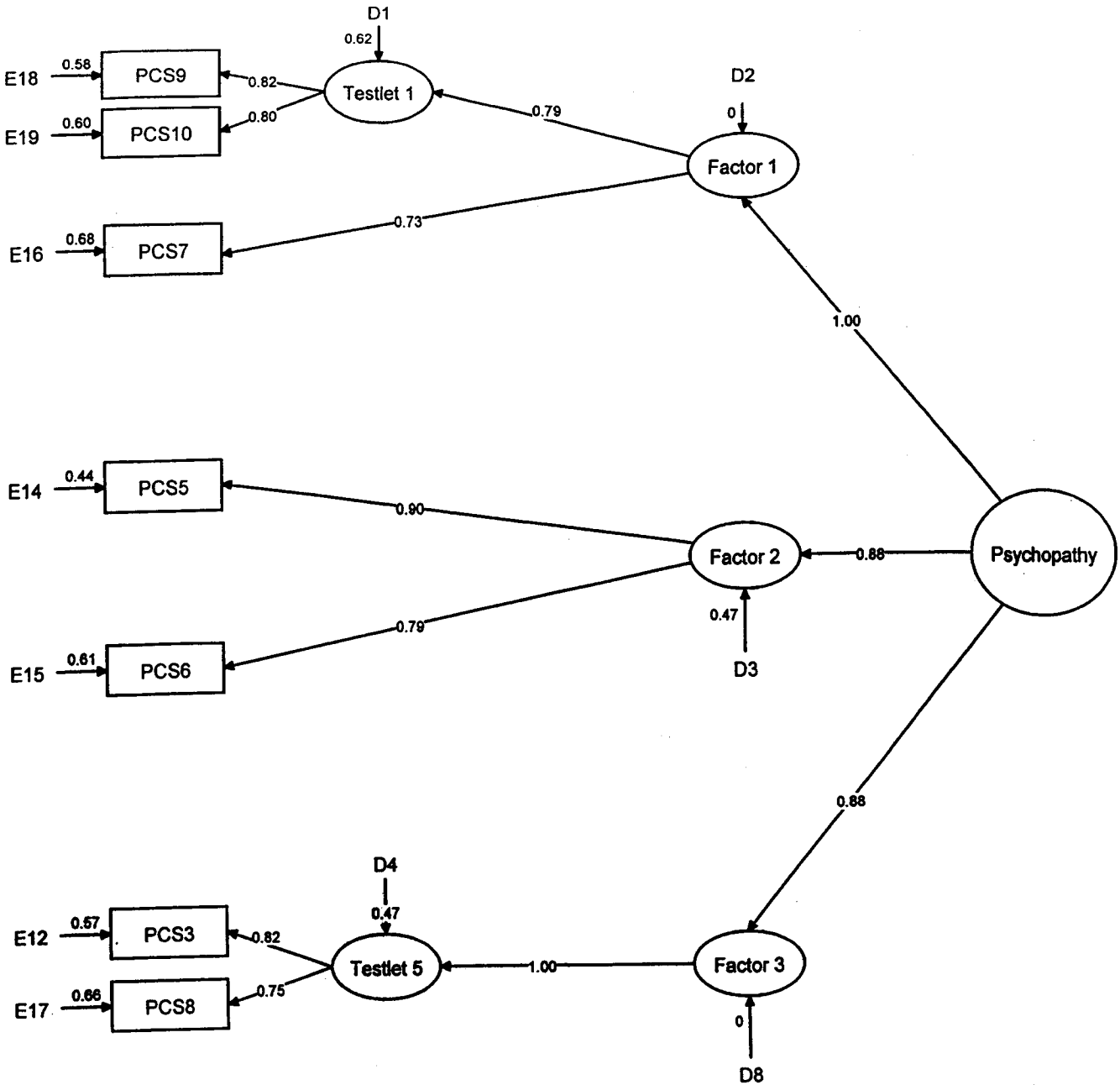

Figure 4. Standardized estimates for the hierarchical factor model derived by confirmatory factor analysis of Psychopathy Criterion Set (PCS) items.

a fit; however, compared with previous models, the fit was poor, $\chi^2(11, N = 506) = 151.9$, $p < .001$, AIC = 129.9, CAIC = 72.4, NFI = .89, NNFI = .80, CFI = .90, GFI = .91, AGFI = .77, RMSEA = .16. This finding was not surprising given that only one item specified the Arrogant and Deceitful Interpersonal Style factor and that the Impulsive and Irresponsible Behavioral Style factor was specified by only one testlet.

We then decided to develop a model using variables derived from all three criteria sets. Using the PCS model as the base model, we considered the local dependence among these PCS criteria and ASPD and DPD criteria and derived the model displayed in Figure 6.

Results

The Arrogant and Deceitful Interpersonal Style factor was underpinned by two testlets, one that combined the PCS items glib and superficial with inflated and arrogant self-appraisal and one that combined the PCS item deceitful and manipulative and the *DSM–IV* item no regard for the truth. The Deficient Affective Experience factor was underpinned by two testlests. The first testlet contained the PCS item lacks remorse, the *DSM–IV* item lacks remorse, and two ICD–10 items, incapacity for guilt and

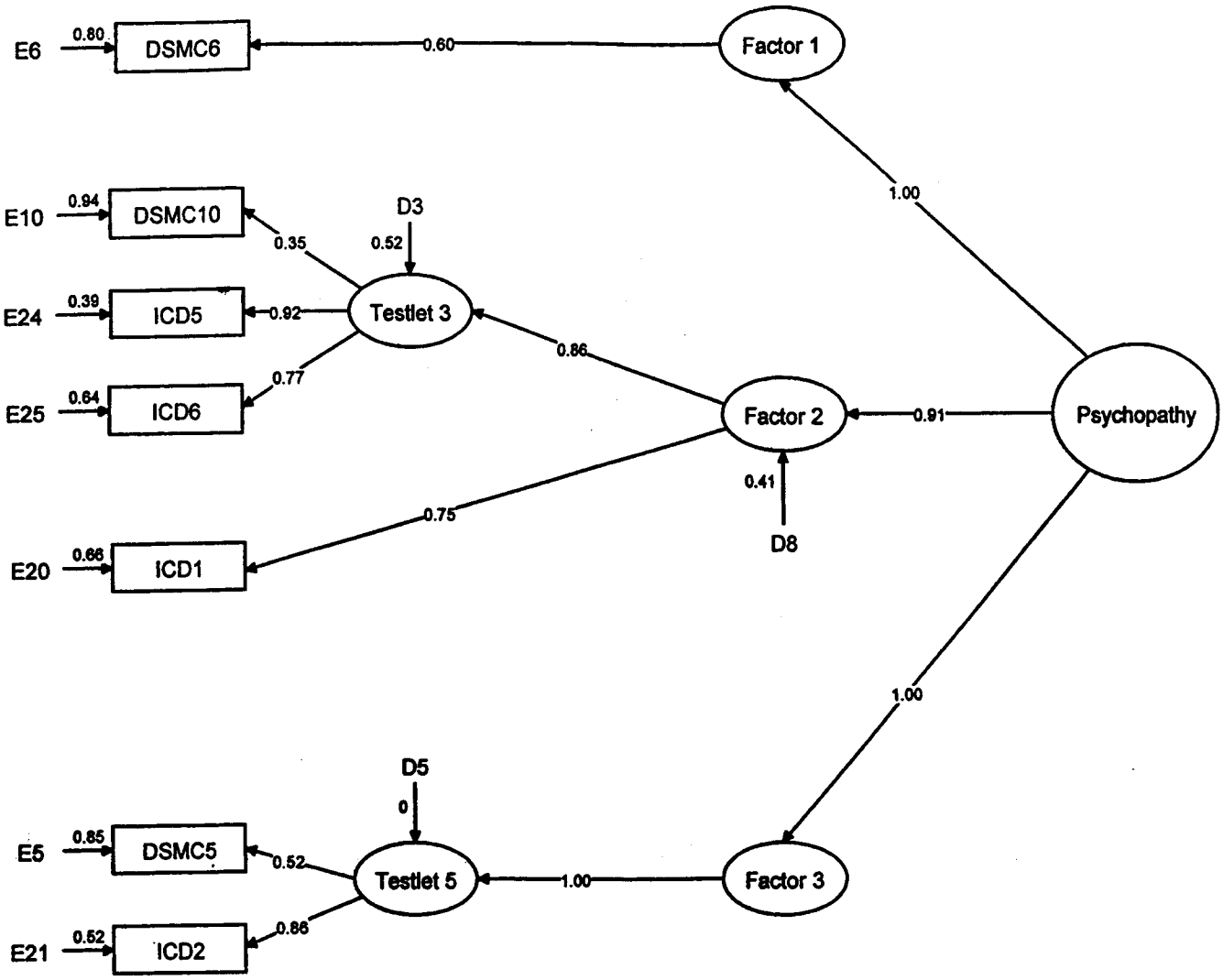

Figure 5. Standardized estimates for the hierarchical factor model derived by confirmatory factor analysis of items derived from the *Diagnostic and Statistical Manual of Mental Disorders,* third edition, revised (DSM), and ICD–10 criteria sets.

profit from experience and proneness to rationalize and blame others. The second testlet contained the PCS item lacks empathy and the ICD–10 item callous unconcern and lack of empathy. The Impulsive and Irresponsible Behavioral Style factor was underpinned by one testlet composed of the PCS items impulsive and irresponsible, the *DSM–IV* item fails to plan ahead or impulsive, and the ICD–10 item persistent irresponsibility and norm disregard. These three factors loaded strongly on the superordinate factor (arithmetic mean loading = .95). The fit indices indicates that this model provides a good fit to the data, $\chi^2(69, N = 486) = 215.6$, $p < .001$, AIC = 77.6, CAIC = −280.3, NFI = .96, NNFI = .96, CFI = .97, GFI = .94, AGFI = .91, RMSEA = .07. The estimated GFS indicates that the superordinate factor in this model is essentially a coherent construct (GFS = .93).

Discussion

This further test of cross-validation provides additional confirmation of the generalizability of the model. However, it is noteworthy that neither the ASPD nor the DPD criteria provides items that adequately represent the Deceitful Interpersonal Style factor.

General Discussion

In this discussion, we consider the two overarching themes raised in the introduction. First, is there evidence for a coherent syndrome of psychopathy? Second, what are the core features of psychopathy? We begin by considering the status of the two-factor model.

Earlier, we argued that the two-factor model is founded on ad hoc statistical methods, in particular, the misinterpretation of congruence coefficients. Reanalysis of the data using more powerful and more appropriate methods confirms that the two-factor model does not provide an adequate description of psychopathy. Three first-order factors appear to be necessary to specify the superordinate construct of psychopathy.

The hierarchical three-factor model developed using PCL–R data is robust: It cross-validated within North America and across

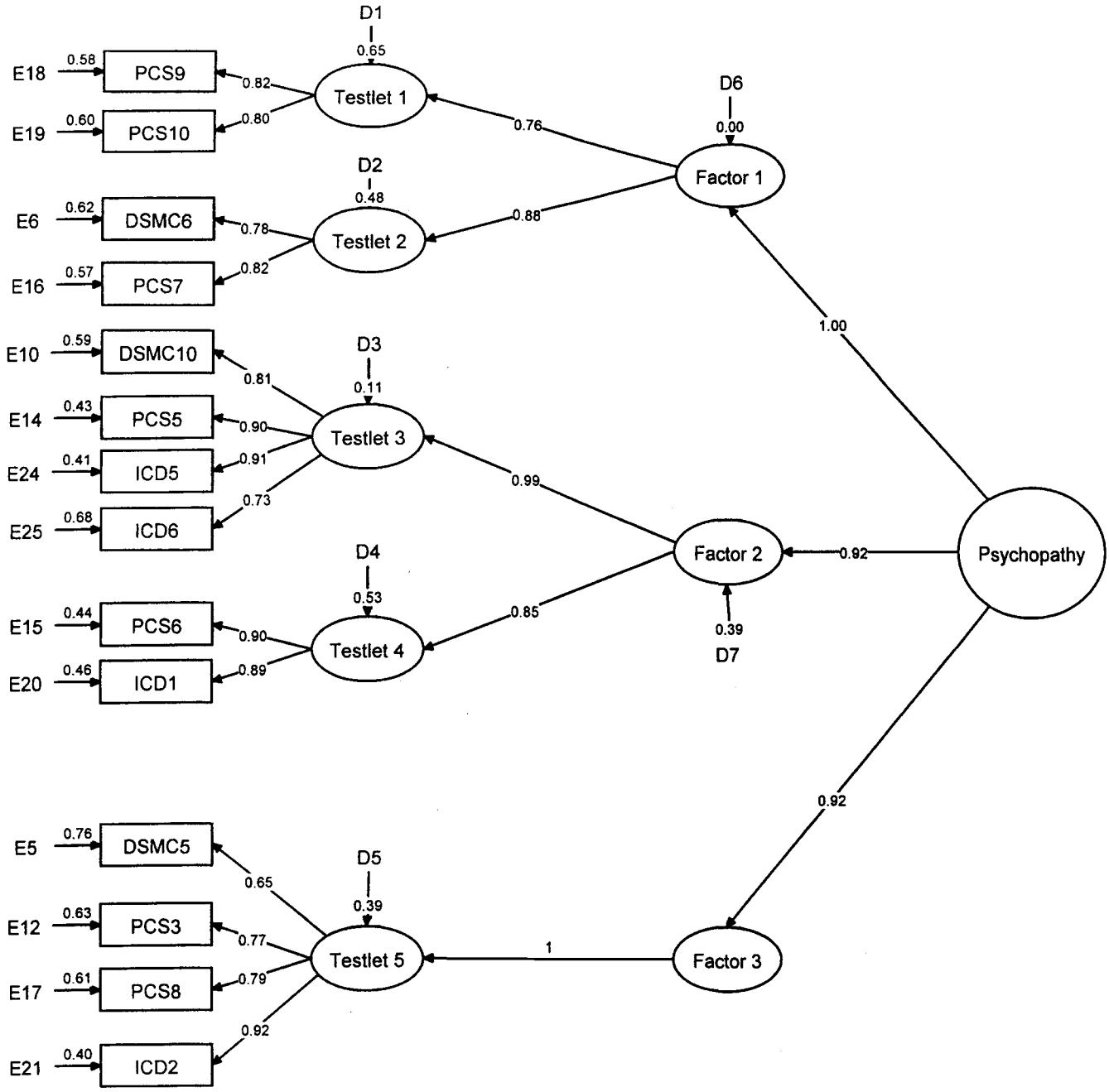

Figure 6. Standardized estimates for the hierarchical factor model derived by confirmatory factor analysis of items derived from the Psychopathy Criterion Set (PCS); *Diagnostic and Statistical Manual of Mental Disorders* (DSM), third edition, revised; and ICD-10 criteria sets.

cognate measures of psychopathy, namely, the PCL:SV, the PCS, and items derived from the ASPD and ICD–10 criteria. Cross-validation is an important indicator of the adequacy of a model (Byrne, 1994; MacCallum et al., 1992; Van de Vijver & Leung, 1997).

Although the model provides a just adequate fit to the Scottish data, some lack of fit is associated with the Arrogant and Deceptive Interpersonal Style factor. This is consistent with previous IRT analysis (Cooke & Michie, 1999) on a subset of the Scottish data analyzed in this article. The interpersonal features, in particular, grandiose sense of self-worth and glibness/superficial charm, demonstrated cross-cultural bias. IRT analysis indicates that Scottish participants have to be significantly higher on the underlying trait of psychopathy than North American participants before they score positively on these items. This current analysis lends support to these earlier findings.

Evidence for a Coherent Construct

As noted earlier, a fundamental prerequisite for a valid construct of psychopathy is the demonstration of a coherent syndrome (Blashfield & Draguns, 1976; Eysenck, 1970; E. Robins & Guze, 1970). The new model indicates that three distinct facets are subsumed by a higher order factor. Critically, to provide evidence for a coherent syndrome, it is not sufficient merely to demonstrate the presence of a higher order factor; it is also necessary to demonstrate that the higher order factor is saturated (Zinbarg et al., 1997). The factor saturation for the higher order factor derived from the North American PCL–R data is .77; higher values were achieved with the other data sets. This indicates that the superordinate factors in the models represent coherent psychological measures (Cronbach, 1951).

The demonstration of a saturated higher order factor has importance for our understanding of the nature of psychopathy. Referring to the two-factor model, Lilienfeld (1994) posed the question "What is psychopathy?" (p. 28)—he asked which of the two traditional factors represent the core of the disorder. Others have raised similar questions (see, e.g., Lilienfeld, 1998; Salekin et al., 1996; Widiger & Lynam, 1998). Our analysis indicates that all three factors are necessary for the characterization of the disorder—each factor contributing to the superordinate factor to a similar extent.

From a practical perspective, this coherence confirms that it is legitimate to sum item scores to provide a general measure of psychopathy. Additional understanding may be gained by estimating scores for the three individual factors just as, for example, different information may be gained from the Full, Verbal, and Performance scores of the Wechsler Adult Intelligence Scale (Zinbarg & McDonald, 1999). For example, Tiihonen and his colleagues demonstrated that the Arrogant and Deceptive Interpersonal Style factor correlates significantly more highly with right amygdala volume than with either the total PCL–R score or the old Factor 1 (J. Tiihonen, personal communication, May 22, 2000).

Personality Construct or Behavioral Construct?

The new model places the definition of psychopathy firmly within the domain of personality pathology. Blackburn (1988, 1998) argued that conceptual confusion emerges with the failure to distinguish between the domains of personality disposition and antisocial behavior. Antisocial behavior may arise from a multiplicity of causes of which personality pathology is only one. Conflating traits and behavior in the measurement of psychopathy means that it is impossible to infer that personality pathology drives antisocial behavior.

Blackburn's (1988) distinction resonates with McCrae and Costa's (1995) distinction between *basic tendencies* and *characteristic adaptations,* basic tendencies being core personality traits and characteristic adaptations being the product of the interaction between the basic tendencies and sociocultural influences. Lilienfeld (1994, 1998) has drawn attention to the importance of this distinction in relation to the PCL–R. He argued that certain PCL–R items represent underlying personality traits that can be expressed in a variety of ways, whereas other PCL–R items represent outward behavioral manifestations that can reflect combinations of personality traits (S. O. Lilienfeld, personal communication, February 1, 1999). Clearly, this is not an all-or-none distinction; it is a matter of degree. We consider that the new model contains items that capture basic tendencies rather than characteristic adaptations. This shift in emphasis has implications for the dominant models in the field.

Two conceptual traditions are apparent in the literature (e.g., Hart & Hare, 1997; Lilienfeld, 1994), one derived from the European and North American construct of psychopathy as exemplified by the diagnosis of DPD (World Health Organization, 1992) and one emerging from the neo-Kraeplinian movement in psychodiagnosis (see, e.g., L. N. Robins, 1966). The first approach is essentially personality based, whereas the second emphasizes publicly observable behaviors. Our new model sharpens this distinction, removing the emphasis on specific behaviors and shifting the emphasis strongly toward the personality domain.

Criminality: Core Feature, Correlate, or Consequence?

The new model places little emphasis on criminality; six of the seven deleted items explicitly or implicitly entail criminal behavior. The emphasis on criminal behavior in the PCL–R may in part reflect the population in which the instrument was developed. Our model implies that though criminality is not a core feature of psychopathy, criminality is certainly a correlate and may even be a consequence (Hart & Hare, 1997). This lesser emphasis on criminality in the new model is consonant with the clinical tradition. Schneider (1950/1958) argued that psychopaths are well represented in society beyond criminal groups; indeed, he argued that they could be unusually successful in positions of political power. Many psychopaths have no antisocial history, and the converse is equally true, many individuals with chronic antisocial behavior not being psychopaths (Cleckley, 1976; Hare, 1993; Lykken, 1995; Widom, 1977).

Harris, Rice, and Quinsey (1994) argued that psychopathy is a taxon, that is, a natural category or nonarbitrary class. The majority of the items that they used to define the taxon have been dropped from the three-factor model. This suggests that the taxon, if it exists, is not psychopathy but may represent the life-course persistent offender taxon described by Moffitt (1993).

Future Directions

Revising measures of psychopathy. Having extracted a coherent construct of psychopathy, it is possible to assess the impact of modifying current items and writing new items on the assessment of the disorder (Steinberg & Thissen, 1996). This process has yielded dividends before. In the development of the PCL:SV, some items were combined, some were simplified, and the amount of collateral information required was reduced. The new items provided either the same or more information about the underlying trait (Cooke et al., 1999). Disaggregating complex PCL–R items into their component parts may improve precision of measurement. For example, the PCL–R item need for stimulation/proneness for boredom could be disaggregated into its component parts, one relating to the excessive need for stimulation and the other relating to low self-discipline, in particular, the ability to complete tasks despite boredom. This process should clarify the critical aspect (or aspects) of this item.

Two other forms of modification should improve both the practical process of completing ratings and the precision of mea-

surement. First, items that form testlets should be deconstructed and rewritten as distinct items. Second, new items designed to measure more specific personality facets that have been implicated in the description of psychopathy, for example, emotional coldness, incapacity for love, egocentricity, fearlessness, and absence of anxiety, should be developed (Harris et al., 1994; Poythress, Edens, & Lilienfeld, 1998). The performance of modified items can then be compared with the performance of the original items and retained or discarded dependent on empirical results (Van der Linden & Hambleton, 1996).

We are advocating the future revision of the full PCL–R to assist research into the nature of the disorder; however, we strongly emphasize the necessity of continuing to use the full PCL–R for risk assessment and other applied purposes.

Validation of the new model. Applying CFA to specify the phenotypic structure of a disorder is merely the starting point, not the end point, of research. The validity and utility of the new model must be assessed. It is necessary to explore the nomological network linking these factors to key variables, for example, developmental, clinical, criminological, and psychophysiological variables. The two-factor model demonstrates differential associations with a wide range of variables (Hare et al., 1990). Fortunately, because of the pervasiveness of the PCL–R in clinical and laboratory research, it will be possible to reanalyze data to determine how the nomological network can be clarified by the new model.

From a clinical perspective, there is a need to discover whether refining the structure improves understanding about the impact of psychopathy on clinically important variables including general and violent recidivism and treatment outcome. For example, can Hart and Dempster's (1997) observation of a link between the old Factor 1 and planned predatory violence be explained by the deficits in empathy, fear, and guilt that form the Deficient Affective Experience factor rather than by the characteristics that underpin the Arrogant and Deceitful Interpersonal Style? Are the disruptive therapeutic group behavior and poor therapeutic outcome that have been observed (see, e.g., Hobson et al., 2000; Hughes et al., 1997; Seto & Barbaree, 1999) more closely linked to interpersonal style than to the lack of appropriate emotions? Clarifying these associations would have clinical relevance for the treatment of this important disorder (Lösel, 1998).

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barrett, P. (1986). Factor comparison: An examination of three methods. *Personality and Individual Differences, 7,* 327–340.

Bentler, P. M., & Wu, E. J. C. (1995). *EQS for Windows.* Encino, CA: Multivarate Software.

Blackburn, R. (1988). On moral judgements and personality disorders: The myth of psychopathic personality revisited. *British Journal of Psychiatry, 153,* 505–512.

Blackburn, R. (1998). Psychopathy and personality disorder: Implications of interpersonal theory. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 269–301). Dordrecht, the Netherlands: Kluwer Academic.

Blashfield, R. K., & Draguns, J. G. (1976). Evaluative criteria for psychiatric classification. *Journal of Abnormal Psychology, 85,* 140–150.

Blashfield, R. K., & McElroy, R. A. (1995). Confusions in the terminology used for classificatory models. In J. E. Livesley (Ed.), *The DSM–IV personality disorders* (pp. 407–416). New York: Guilford Press.

Brandt, J. R., Kennedy, W. A., Patrick, C. J., & Curtin, J. J. (1997). Assessment of psychopathy in a population of incarcerated adolescent offenders. *Psychological Assessment, 9,* 429–435.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows.* London: Sage.

Chen, W. H., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22,* 265–289.

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St Louis, MO: Mosby.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooke, D. J. (1995). Psychopathic disturbance in the Scottish prison population: The cross-cultural generalisability of the Hare Psychopathy Checklist. *Psychology, Crime and Law, 2,* 101–108.

Cooke, D. J., Forth, A., & Hare, R. D. (1998). *Psychopathy: Theory, research and implications for society.* Dordrecht, The Netherlands: Kluwer Academic.

Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist—Revised. *Psychological Assessment, 9,* 3–14.

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108,* 55–68.

Cooke, D. J., Michie, C., Hart, S. D., & Hare, R. D. (1999). The functioning of the Screening Version of the Psychopathy Checklist—Revised: An item response theory analysis. *Psychological Assessment, 11,* 3–13.

Cooke, D. J., Michie, C., Philip, L., & Carr, E. (1997). *Dimensions of psychopathy and risk for instrumental and reactive violence.* Paper presented at the 7th annual conference of the Division of Criminological and Legal Psychology of the British Psychological Society, Leciester, England.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653–665.

Côté, G., & Hodgins, S. (1989). *Psychopathy Checklist: Validation of a French version.* Unpublished manuscript.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Darke, S., Kaye, S., Finlay-Jones, R., & Hall, W. (1998). Factor structure of psychopathy among methodone maintenance patients. *Journal of Personality Disorders, 12,* 162–171.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research, 18,* 197–218.

Everitt, B. S. (1993). *Cluster analysis.* New York: Arnold.

Eysenck, H. J. (1947). *Dimensions of personality.* London: Routledge & Kegan Paul.

Eysenck, H. J. (1970). The classification of depressive illness. *British Journal of Psychiatry, 117,* 241–250.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286–299.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hare, R. D. (1978). Electrodermal and cardiovascular correlates of psychopathy. In R. D. Hare & D. Schalling (Eds.), *Psychopathic disorder: Approaches to research* (pp. 107–143). Chichester, England: Wiley.

Hare, R. D. (1991). *Manual for the Revised Psychopathy Checklist* (1st ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D. (1993). *Without conscience: The disturbing world of the psychopaths among us* (1st ed.). New York: Pocket Books.

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23,* 25–34.

Hare, R. D., Cooke, D. J., & Hart, S. D. (1999). Psychopathy and sadistic personality disorder. In T. Millon & P. H. Blaney (Eds.), *Oxford textbook of psychopathology.* Oxford, England: Oxford University Press.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The Revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 238–241.

Harpur, T. J., Hakstian, A., & 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., & 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.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology, 62,* 387–397.

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version* (1st ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hart, S. D., & Dempster, R. J. (1997). Impulsivity and psychopathy. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 212–232). New York: Guilford Press.

Hart, S. D., & Hare, R. D. (1997). Psychopathy: Assessment and association with criminal conduct. In D. M. Stoff, J. Maser, & J. Brieling (Eds.), *Handbook of antisocial behavior* (pp. 22–35). New York: Wiley.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3,* 139–170.

Hobson, J., & Shine, J. (1998). Measurement of psychopathy in a UK prison population referred for long-term psychotherapy. *British Journal of Criminology, 38,* 504–515.

Hobson, J., Shine, J., & Roberts, R. (2000). How do psychopaths behave in a therapeutic community? *Psychology, Crime and Law, 6,* 139–154.

Hughes, G., Hogue, T. E., Hollin, C. R., & Champion, H. (1997). First stage evaluation of a treatment programme for personality disordered offenders. *Journal of Forensic Psychiatry, 8,* 515–527.

Intrator, J., Hare, R. D., Stritzke, P., Brichtswein, K., Dorfman, D., Harpur, T., Berstein, D., Handelsman, L., Schaefer, C., Keilp, J., Rosen, J., & Machac, J. (1995). A brain imaging (SPECT) study of semantic and affective processing in psychopaths. *Biological Psychiatry, 42,* 96–103.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Kosson, D. S., Smith, S. S., & Newman, J. P. (1990). Evaluating the construct validity of psychopathy in Black and White male inmates: Three preliminary studies. *Journal of Abnormal Psychology, 99,* 250–259.

Lapierre, D., Braun, C. M. J., Hodgkins, S., & Toupin, J. (1995). Neuropsychological correlates of violence in schizophrenia. *Schizophrenia Bulletin, 21,* 253–262.

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

Lilienfeld, S. O. (1998). Methodological advances and developments in the assessment of psychopathy. *Behaviour Research and Therapy, 36,* 99–125.

Livesley, W. J. (1998). The phenotypic and genotypic structure of psychopathic traits. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 69–81). Dordrect, The Netherlands: Kluwer Academic.

Livesley, W. J., Jackson, D. N., & Schroeder, M. (1992). Factorial structure of traits delineating personality disorders in clinical and general population samples. *Journal of Abnormal Psychology, 101,* 432–440.

Livesley, W. J., & Schroeder, M. (1991). Dimensions of personality disorder: The *DSM–III–R* cluster B diagnosis. *Journal of Nervous and Mental Disease, 179,* 320–328.

Lösel, F. (1998). Treatment and management of psychopaths. In D. J. Cooke, A. Forth, & R. D. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 303–354). Dordrecht, The Netherlands: Kluwer Academic.

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

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model specifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490–504.

Marshall, L. A., & Cooke, D. J. (1999). The childhood experiences of psychopaths: A retrospective study of familial and societal factors. *Journal of Personality Disorders, 13,* 211–225.

McCrae, R. R., & Costa, P. T. (1995). Trait explanations in personality psychology. *European Journal of Personality, 9,* 231–252.

McDermott, P. A., Alterman, A. I., Cacciola, J. S., Rutherford, M. J., Newman, J. P., & Mulholland, E. M. (2000). Generality of Psychopathy Checklist—Revised factors over prisoners and substance-dependent patients. *Journal of Consulting and Clinical Psychology, 68,* 181–186.

Michie, C., & Cooke, D. J. (2000). *The structure of violent behavior: A hierarchical model of the MacArthur Community Violence Screening Instrument.* Manuscript submitted for publication.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Newman, J. P., & Wallace, J. F. (1993). Psychopathy and cognition. In P. Kendell & K. Dobson (Eds.), *Psychopathology and cognition* (pp. 293–349). New York: Academic Press.

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. *Journal of Abnormal Psychology, 103,* 82–92.

Pham, T. H. (1998). Psychometric evaluation of Hare's Psychopathy Checklist—Revised amongst a population of incarcerated Belgian prisoners. *L'Encephale, 24,* 435–441.

Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (1998). Criterion-related validity of the psychopathic personality inventory in a prison sample. *Psychological Assessment, 10,* 426–430.

Rice, M. E., & Harris, G. T. (1992). A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders. *International Journal of Law and Psychiatry, 15,* 397–408.

Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry, 126,* 983–987.

Robins, L. N. (1966). *Deviant children grown up: A sociological and psychiatric study of sociopathic personality.* Baltimore: Williams & Wilkins.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist—Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice, 3,* 203–215.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1997). Construct validity of psychopathy in a female offender sample: A multitrait–multimethod evaluation. *Journal of Abnormal Psychology, 106,* 576–585.

Schneider, K. (1958). Psychopathic personalities (9th ed.; M. Hamilton, Trans.). London: Cassell. (Original work published 1950)

Serin, R. C. (1996). Violent recidivism in criminal psychopaths. *Law and Human Behavior, 20,* 207–217.

Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence, 14,* 1235–1248.

SPSS. (1993). *SPSS users' guide.* New York: McGraw-Hill.

Steinberg, L., & Thissen, D. (1996). Uses of item response theory and the

testlet concept in the measurement of psychopathology. *Psychological Measurement, 1,* 81–97.

Ten Berge, J. M. F. (1986). Rotation to perfect congruence and cross-validation of component weights across populations. *Multivariate Behavioural Research, 21,* 41–64.

Van de Vijver, F. J., & Leung, K. (1997). *Methods and data analysis for cross-cultural research.* Thousand Oaks, CA: Sage.

Van de Vijver, F. J., & Poortinga, Y. H. (1994). Methodological issues in cross-cultural studies on parental rearing behavior and psychopathology. In C. Perris, W. A. Arrindell, & E. Eisemann (Eds.), *Parental rearing and psychopathology* (pp. 173–197). Chichester, England: Wiley.

Van der Linden, W. J., & Hambleton, R. K. (1996). *Handbook of modern item response theory.* New York: Springer.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of Abnormal Psychology, 103,* 18–31.

Widiger, T. A., Cadoret, R., Hare, R. D., Robins, L. N., Rutherford, M. J., Zanarini, M., Alterman, A. I., Apple, M., Corbitt, E., Forth, A. E., Hart, S. D., Kultermann, J., & Woody, G. (1996). *DSM–IV* antisocial personality disorder field trial. *Journal of Abnormal Psychology, 105,* 3–16.

Widiger, T. A. E., & Lynam, D. R. (1998). Psychopathy and the five-factor model of personality. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davies (Eds.), *Psychopathy: Antisocial, criminal and violent behavior* (pp. 171–187). New York: Guilford Press.

Widom, C. S. (1977). A methodology for studying noninstitutionalized psychopaths. *Journal of Consulting and Clinical Psychology, 45,* 674–683.

World Health Organization. (1992). *The ICD–10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines* (1st ed.). Geneva, Switzerland: Author.

Zinbarg, R. E., & Barlow, D. H. (1996). The structure of anxiety and anxiety disorders: A hierarchical model. *Journal of Abnormal Psychology, 105,* 181–193.

Zinbarg, R. E., Barlow, D. H., & Brown, T. A. (1997). Hierarchical structure and general factor saturation of the Anxiety Sensitivity Index: Evidence and implications. *Psychological Assessment, 9,* 277–284.

Zinbarg, R. E., & McDonald, R. D. (1999). *General factor saturation and the internal structure of tests.* Unpublished manuscript.

Received February 7, 2000
Revision received December 5, 2000
Accepted February 1, 2001 ■