Psychopathy Across Cultures: North America and Scotland Compared

David J. Cooke
Glasgow Caledonian University and Douglas Inch Centre

Christine Michie
Glasgow Caledonian University

Differences in the prevalence and presentation of psychopathic personality disorder between North America and Scotland were evaluated. R. D. Hare's (1991) Psychopathy Checklist—Revised ratings obtained from a sample of 2,067 North American male prisoners and forensic patients were compared with ratings obtained from 246 Scottish male prisoners. Item response theory methods were used to examine differences in the performance of items and to equate the scale across settings. The items had equal relevance to the description of psychopathic personality disorder in both settings; however, the Scottish prisoners had to have higher levels of the underlying latent trait before certain characteristics became apparent. The prevalence of the disorder appears to be lower in Scotland. Explanations for the observed differences in terms of enculturation, socialization, and migration are explored.

Within North America, psychopathic personality disorder is a syndrome characterized by specific behavioral, affective, and interpersonal characteristics (Cleckley, 1976; Hare, 1991). Behaviorally, psychopaths are impulsive and sensation seeking, they violate social norms and frequently become involved in criminal activity. Affectively, psychopaths are emotionally labile and shallow; they do not experience empathy, guilt, or remorse. Interpersonally, psychopaths are manipulative, grandiose, egocentric, forceful, and cold.

Psychopaths appear to exist across time and across cultures (e.g., Cleckley, 1976; Cooke, 1996, 1998). Referring to the cognate construct of Antisocial Personality Disorder (ASPD), Robins, Tipp, and Przybeck (1991) argued that the disorder is recognized in all societies, irrespective of their level of economic development or their era. Robins et al. (1991) emphasized, however, that the prevalence of ASPD varies with time and place. Murphy's (1976) research has emphasized the ubiquity of the disorder: She demonstrated that groups as distinct as the Inuit of northwestern Alaska and the Yoruba of Nigeria have a concept of psychopathy. They can distinguish psychopathy from other mental disorders: Psychopathy is rare in these settings.

The dominant North American models of personality and psychopathology may not take into account the extensive cross-cultural diversity of psychological phenomena (Fiske, 1995; Lewis-Fernandez & Kleinman, 1994). This may be particularly true for the personality disorders. Compared with mental disorders such as depression and schizophrenia, psychopathy may have a less well-crystallized "pan-cultural core" (Draguns, 1986, p. 333). Fortunately, awareness that diagnostic decisions are not culture free is growing: Explicit instructions to consider cultural features when coming to a diagnosis are provided in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., *DSM–IV*; American Psychiatric Association, 1994).

The nature and expression of the symptoms and signs that characterize the syndrome of psychopathy may be molded by culture: Draguns (1973) contended that psychopathology, in general, may be "an exaggeration or a caricature of the socially shared and prevalent patterns of adaptation" (p. 33). As noted above, this is more likely to be the case for the personality disorders than for Axis I disorders such as schizophrenia or depressive disorders. Weisz, Suwanlert, Chaiyasit, and Walter (1987), discussing their findings on differences in the prevalence of under- and overcontrolled behavior in children, argued for a suppression-facilitation model. They indicated that "characteristics of a culture—values, beliefs, expectancies, and child-rearing practices—may suppress the development of certain types of child behavior problems and foster and facilitate the development of others" (p. 723). Similar processes may also influence psychopathy. Hare (1998) argued that "the behavioral expressions of psychopathy, as well as the degree to which they stand out from the behaviors of others, are influenced by societal and cultural structures and norms" (p. 106). Cultural changes may influence the genesis of this disorder (Paris, 1993). This is not a new idea. In the middle of the last century, Pritchard (cited in Sanchez, 1986) argued that "moral insanity"—a precursor of the diagnosis of psychopathy—was caused by the social changes consequent on industrialization.

In this article, we examine the issue of whether the expression of psychopathy in Scotland is different from the expression of the

disorder in North America. Differences would have both practical and theoretical implications.

Measuring Psychopathy: The Hare (1991) Psychopathy Checklist—Revised (PCL–R)

The PCL–R is currently the instrument of choice for measuring psychopathy (Conoley & Impara, 1995). Dinges, Atlis, and Vincent (1997) argued that the PCL–R is more likely to be sensitive to cross-cultural variation than the diagnosis of APSD: They contend that the "the content of the *ICD-10* [*International Classification of Diseases*, 10th ed.] and PCL–R criteria provides more balanced symptom domains" (p. 465). The PCL–R has high internal consistency and interrater reliability in North America and Scotland (e.g., Cooke, 1989; Hare, 1991).

Within North America, the PCL–R displays notable consistency of psychometric and distributional characteristics across both correctional and forensic samples (Hare, 1991). There have been relatively few studies of the PCL–R outside North America. Cooke (1998) reviewed PCL–R data from 16 European samples, from the Scandinavian countries in the north, through the United Kingdom, Germany, and Belgium, to Spain and Portugal in the south. Comparison of the overall mean in the European sample with the mean of the standardization sample indicated a substantial difference (Cohen's $d = .81$).

The Item Response Theory (IRT) Approach to Cross-Cultural Differences

The cross-cultural generalizability of personality and other psychometric scales has historically been assessed using the procedures of classical test theory (CTT) (e.g., Barrett & Eysenck, 1984; Cooke, 1995b). These CTT approaches are not without their critics; proponents of IRT argue that IRT models may be more appropriate methods for assessing cross-cultural variations in the nature and presentation of a disorder.

We do not provide a detailed account of IRT methods here: The advantages of the methods have been detailed elsewhere (Cooke & Michie, 1997; Embretson, 1996; Steinberg & Thissen, 1996). However, to make the analyses easier to follow, we summarize the salient features of IRT methods in the context of cross-cultural comparisons and endeavor to highlight their advantages.

IRT models are mathematical expressions of the relation between an individual's response or rating on an item and an underlying latent trait or construct that is postulated to underpin these responses or ratings. A mathematical function specifies a trace line—or item characteristic curve (ICC)—that represents the manner in which the probability of a response or score varies with the level of the underlying trait (the underlying trait is represented by the symbol Θ). Θ is standardized with a mean of 0 and a standard deviation of 1. Two characteristics of the curves are important, their slope and their position on the underlying trait. This is illustrated most simply with dichotomous items. The ICCs can be expressed graphically as illustrated in Figure 1.

We have plotted the ICCs for the "2" ("present") response for three PCL–R items derived from North American samples (Cooke & Michie, 1997). The ICC for the PCL–R item "callous/lack of empathy" is steeper than the ICC for the PCL–R items "glibness/superficial charm" and "irresponsibility." This illustrates that the item "callous/lack of empathy" discriminates more effectively between different levels of the trait as compared with the other two items. "Glibness/superficial charm" and "irresponsibility" have equal slopes—they are equally discriminating—but have different positions on the underlying trait. An individual whose psychopathy trait strength is moderate may be irresponsible; by way of contrast, an individual who is glib and superficial is likely to have high psychopathy trait strength. IRT methods have particular advantages in cross-cultural comparisons; these advantages are summarized below.

1. *ICCs are independent of the samples from which they are generated.* Cross-cultural comparisons using traditional CTT methods are based on comparisons of indexes including corrected item-to-total correlations and Cronbach's alpha; these indexes are highly sensitive to variations in the range of test scores across samples. ICCs are independent, within a linear transformation, of the samples from which they are derived (Mellenbergh, 1996).

2. *Representative samples are not required.* Embretson (1996) indicated that IRT methods mean that many of the "old rules" of psychometrics no longer apply. One critical "rule change" from the perspective of cross-cultural or cross-group comparisons is that it is not necessary to have representative samples to obtain unbiased estimates of item properties; unbiased estimates can be derived from nonrepresentative samples (Embretson, 1996).
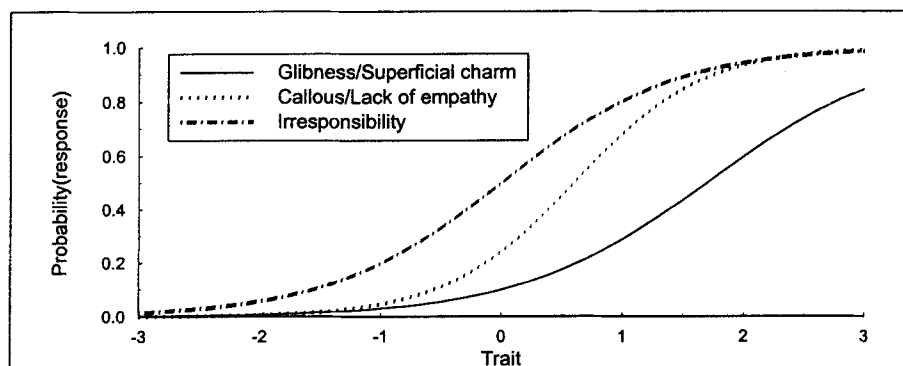

Figure 1. Illustrative item characteristic curves for three Psychopathy Checklist—Revised items derived from North American samples. Curves are for a "2" response on each item.

3. *IRT analyses permit direct comparison of parallel items.* That is, it is possible to compare directly items measuring different characteristics in the same sample or the same item in different samples. This feature of the methodology means that IRT methods can be used effectively to determine whether item and test scores are invariant across forms (e.g., original vs. revision or translation; full vs. short form) and across respondents (e.g., men vs. women; young vs. old). For example, IRT methods have been used to demonstrate that the Screening Version of the PCL–R measures the same underlying trait as the full PCL–R (Cooke, Michie, Hart, & Hare, in press).

4. *IRT procedures can be used to identify items that perform differently across cultures.* Differential item functioning (DIF) occurs when individuals from different cultures with the same level of ability or the same trait strength (i.e., the same value of Θ) obtain different scores on an item. Differences in scores may reflect differences in the relevance of the item across cultures or differences in the language used to define the items across cultures. IRT methods have utility in detecting DIF across cultures (Bontempo, 1993; Holland & Wainer, 1993; Reise, Widaman, & Pugh, 1993).

5. *IRT procedures can ensure measurement invariance across groups.* To compare estimates of the prevalence of psychopathy in two different settings, it is necessary not only to establish that the same trait is being measured but also to establish that it is measured on the same measurement scale (Reise et al., 1993). The Fahrenheit and Celsius scales measure the same construct, but both of their zero points and the size of their intervals are different. Clearly, if a scale does not display measurement invariance across groups, then cross-group comparisons of prevalence are essentially meaningless. It is not possible to determine whether measurement invariance exists with CTT; procedures for this purpose do exist in IRT (Reise et al., 1993). Even if items on the scale behave differently across samples, as long as there is a core group of items that behave the same way in both samples, it is possible to generate a common metric (i.e., a common scale for Θ) on which to measure the trait in both groups (Reise et al., 1993).

It is important for clarity to emphasize the distinction between individuals' scores on the underlying trait as measured by the IRT model and their scores on the test—in this case the PCL–R. "Anchoring" procedures ensure that the metric of the underlying trait (Θ) is the same; thus an individual in Scotland whose score on Θ is 1.5 will be equivalent in terms of degree of psychopathy to a North American individual whose score on Θ is 1.5. However, these two individuals may have different PCL–R scores; the relation between the underlying common trait metric (Θ) and the scores on the PCL–R may differ across cultures.

The Present Study

The overall purpose of the present study was to determine the extent to which the North American conceptualization of psychopathy—as operationalized using the PCL–R—can be generalized to Scotland. Is a Scottish psychopath the same as a North American psychopath; do the characteristics specified in the PCL–R have relevance for the description of psychopathy in Scotland; is there evidence that is consistent with the suppression or facilitation of particular features of psychopathy in Scotland; are there as many psychopaths in Scotland as in North America?

Method

Participants

The North American samples. Data from 10 North American samples, 8 Canadian and 2 from the United States, were obtained. These samples were essentially convenience samples collected for a range of clinical and experimental purposes. The Canadian samples included 4 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; 65 patients of a forensic outpatient clinic in Vancouver, British Columbia) and 4 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 composed the standardization samples used in the PCL–R manual (Hare, 1991). In addition, data from adult male prisoners in a minimum security institution in Wisconsin were obtained; these data consisted of a sample of 838 White prisoners and 153 Black prisoners (see Cooke & Michie, 1997, for more details).

The Scottish sample. Cooke (1994, 1995b) collected data on a systematic random sample of 307 prisoners in the Scottish prison system; unlike the North American samples, this sample was unique in that it was designed to provide a truly representative sample of a country's prison population. The Scottish prison system is a unified system holding all prisoners either sentenced or on remand. The sample was representative of all those incarcerated in Scotland: Sentence lengths ranged from 7 days to life imprisonment, 19% were female ($n = 61$), the modal age was in the range of 16–20 years, and three quarters of the sample were unemployed at the point of arrest. Eighteen percent of the male participants were living in unsettled accommodation or psychiatric hospitals or sleeping on the street immediately prior to their imprisonment. (This sample has been described in greater detail elsewhere; Cooke, 1994, 1995a, 1995b.)

Materials

The PCL–R is composed of 20 items; a trained rater uses explicit and specific definitions for each 3-point item. The rater determines how closely the individual participant meets the characteristics specified in the item descriptors. Each item is scored on a 3-point scale, 0 (*absent*), 1 (*maybe/in some respects*), or 2 (*present*), indicating the degree to which the item can be said to apply to the individual participant. Information is collected by interview and file review. The items include the behavioral, affective, and interpersonal items thought to characterize psychopathic personality disorder (Cleckley, 1976; Hare, 1991). The 20 items are listed in Table 1.

The PCL–R requires raters to make judgments about whether individual participants display specific characteristics; these characteristics are defined in considerable detail within the test manual (Hare, 1991). PCL–R assessments in the Scottish sample were carried out by two trained evaluators (including David J. Cooke, who is an authorized PCL–R trainer) on the basis of a lengthy interview and a file review. Previous analyses demonstrated that the ratings obtained had psychometric properties similar to the standardization sample (Cooke, 1995b). The instrument is regarded as measuring a superordinate construct underpinned by two correlated lower order factors (Hare, 1991). Ratings for the Scottish sample displayed a very similar pattern to that obtained in the standardization sample (e.g., factor congruence for Factor 1 $= .92$ and Factor 2 $= .93$; correlations between the two factors, $r = .55$). Alpha coefficients for the total score, Factor 1, and Factor 2 were comparable with those obtained in the standardization sample (total score, $\alpha = .88$; Factor 1, $\alpha = .73$; Factor 2, $\alpha = .79$). Overall, these results suggest that the underlying structural properties of the overall test generalize across cultures (van de Vijver & Leung, 1997). These Scottish data were compared with the North American data to provide a cross-cultural comparison at the item level.

Table 1

Item Parameters for North America (NA) and Scotland (SC); Each Item Is Distinct (Full Model)

PCL–R item label	a NA	a SC	b_1 NA	b_1 SC	b_2 NA	b_2 SC
1. Glibness/superficial charm	1.4	0.8	−0.1	2.0	1.6	5.8
2. Grandiose sense of self-worth	1.5	0.9	−0.4	1.2	1.3	3.1
3. Need for stimulation	1.6	1.3	−1.3	−0.6	0.2	1.5
4. Pathological lying	1.5	1.1	−0.6	−0.2	1.1	1.6
5. Conning/manipulative	1.5	1.0	−0.4	0.0	1.1	2.3
6. Lack of remorse or guilt	1.7	1.3	−1.5	−0.1	0.0	1.0
7. Shallow affect	1.6	1.2	−0.8	0.2	0.8	1.9
8. Callous/lack of empathy	2.0	1.3	−1.0	0.1	0.6	1.7
9. Parasitic lifestyle	0.9	1.8	−1.4	0.0	1.4	1.1
10. Poor behavioral controls	1.0	1.0	−1.2	−0.7	0.6	1.0
11. Promiscuous sexual behavior	0.8	0.8	−0.8	−1.3	0.8	0.4
12. Early behavior problems	1.0	1.2	−0.3	0.3	1.0	1.1
13. Lack of long-term goals	1.2	1.4	−1.3	−0.3	0.5	0.4
14. Impulsivity	1.4	0.7	−1.8	−0.7	−0.1	1.4
15. Irresponsibility	1.4	1.6	−1.9	−0.4	0.1	1.4
16. Failure to accept responsibility	1.0	0.7	−1.3	−1.1	0.6	2.8
17. Short-term marital relationships	0.7	1.0	0.9	3.2	2.4	4.5
18. Juvenile delinquency	0.7	1.6	−0.7	−0.8	0.5	−0.2
19. Revocation of conditional release	0.7	0.8	−1.4	0.5	−0.1	3.1
20. Criminal versatility	0.9	1.6	−0.4	−0.5	1.6	0.5

Note. The a parameter is a measure of the discriminating power of the item; b_1 and b_2 are the points of infection for a 0 rating and a 2 rating, respectively. PCL–R = Psychopathy Checklist—Revised.

Results

We have three specific objectives for the analyses presented in this article: first, to compare the functioning of individual PCL–R items in North America and Scotland and to determine whether they have similar relevance in both contexts and whether they were subject to DIF; second, to examine the metric equivalence of the scales in both settings; third, to examine the estimated prevalence of the disorder in North American prisons compared with Scottish prisons.

Choice of Model

A range of different IRT models has been developed (Holland & Wainer, 1993). Samejima's graded model is an appropriate model for PCL–R data (Cooke & Michie, 1997; Cooke et al., in press). Under this model, the interrelation between the probability of each possible response to an item and the latent trait can be described by a curve or a trace line. The curves for "0" and "2" responses are symmetric logistic functions. The curve for the "1" response can be found by subtraction because the total probability of all three responses at any level of the trait must be unity. The shape and position of the curves are summarized by the values of three parameters, a, b_1, and b_2 (Thissen, 1991). These curves are illustrated in Figure 2. In contrast to Figure 1, we need three curves to describe all three values of the PCL–R items.

The slopes at the point of inflection for the probability of being given a score of 0 on an item or the probability of being given a score of 2 on an item—$P(0)$ and $P(2)$, respectively—are of the

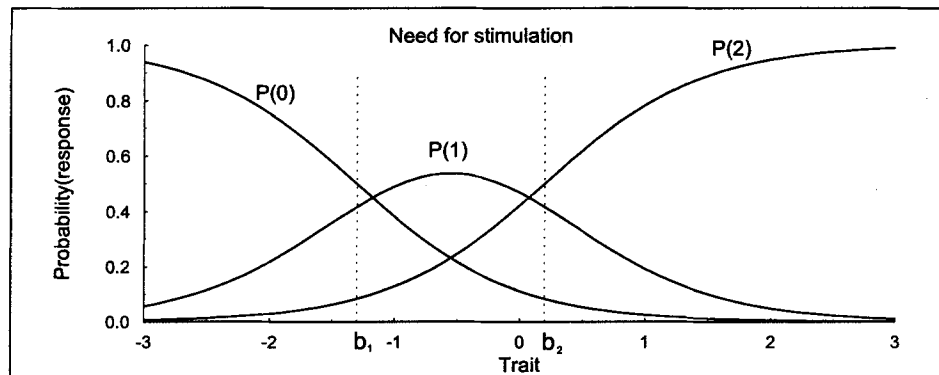
Figure 2. Item response curves for the Psychopathy Checklist—Revised item, "need for stimulation." Curves are for 0, 1, and 2 scores. b_1 and b_2 are parameters indicating the points of infection for a 0 rating and a 2 rating, respectively.

same magnitude but opposite in direction and are determined by the parameter a. This a parameter is a measure of the discriminating power of the item (Holland & Wainer, 1993). The larger the value of a, the steeper the slope. The position of the points of inflection are given by the parameters b_1 for $P(0)$ and b_2 for $P(2)$. At these levels of the trait, the probability of obtaining a score of 0 or a score of 2, respectively, is .5. Thus, parameters b_i provide measures of item difficulty or extremity or frequency of a behavior or attitude. Increases in the value of b_i move the curve to the right, increasing the item's level of extremity, unpopularity, or difficulty.

To ensure that the PCL–R was measuring the same construct in the same way in both Scotland and North America, it was necessary to determine that there was measurement equivalence in the two countries: In IRT terms, item trace lines must be the same across groups. If the item parameters were the same for the two groups, then estimates of the underlying trait across groups should be equivalent. However, if parameters for an item differed across groups, then that item should display DIF. As noted above, the presence of DIF would not be fatal for the determination of cross-group equivalence of measurement; as long as some items have equivalent parameters, equivalent measurement of the underlying trait could still be established.

All IRT analyses were conducted using MULTILOG VI (Thissen, 1991). The program uses the methods of maximum likelihood to estimate the item parameters simultaneously in two or more groups. In line with the standard procedures of generalized likelihood ratio testing (GLRT), the program allows a variety of constraints to be imposed on the parameters. The equivalence of parameters across groups can be determined by comparing the goodness of fit of a constrained model with the the goodness of fit of an unconstrained model. If two models are compared, one in which parameters are constrained to be equivalent across the groups and one in which no such constraints are imposed, and if GLRT reveals no significant difference between the models, this confirms that there is no evidence of any differences in the item parameters across the two groups. Therefore, it is parsimonious to assume no differences. The series of analyses below were based on the methods proposed by Reise et al. (1993).

Analysis of the North American Data

Before investigating the possibility of DIF between Scotland and North America, it was necessary to ensure that no DIF existed among the North American samples. These analyses were described in detail elsewhere and are merely summarized here (Cooke & Michie, 1997). There were 10 North American samples; however, some were small relative to the number of parameters in the IRT model. The data were aggregated into four groups; the Canadian participants were divided into prison participants (n = 636) and forensic participants (n = 440) and the United States participants were divided into White prisoners (n = 838) and Black prisoners (n = 153). In the first step, an unconstrained baseline model was estimated, the item parameters were allowed to vary freely, and, because of variations in sources of the samples (i.e., prisons and forensic settings), the mean levels of psychopathic personality disorder were also allowed to vary in the estimation. In the second step, a new model was estimated in which all item parameters were constrained to be equal across groups; once again, the mean level of trait across groups was allowed to vary.

By comparing the new model with the baseline model, it was possible to determine whether any observed differences in the item parameters across the four groups in the baseline model were merely due to sampling error or whether they were significant differences. Models were compared by determining the difference in the goodness of fit of the two models, as measured by differences in the G^2 statistic ($G^2 = -2 \times$ the log likelihood function). In this case, the increase in G^2 was not statistically significant, indicating no significant DIF among the North American samples.

Comparison of North American and Scottish Data

Differences in item parameters. The first step in the comparison between North America and Scotland was the development of a new baseline model containing the data from both sources. In the baseline model, the North American data were assumed to consist of four groups with different mean levels of psychopathic personality disorder but with the same item parameters. The baseline model was unconstrained because the item parameters and the mean level of psychopathic personality disorder of the Scottish sample were allowed to vary from the North American values. Because the North American data included only male participants, only Scottish men (n = 246) were used in these analyses. The parameters of this baseline model are shown in Table 1.

The presence of DIF between Scotland and North America was determined by constraining the item parameters to be equal in the two settings: The increase in G^2 was statistically highly significant, $G^2(60) = 109$, $p < .001$, indicating that the data could not be modeled adequately if the item parameters were assumed to be the same in the two settings; there was DIF in at least one item.

Inspection of Table 1 suggests that differences in the slope (a) parameters were smaller than those for the threshold (b_1 and b_2) parameters; the standard errors for the Scottish slope parameters were approximately 0.3. The hypothesis that the item slopes were the same in both settings was tested by constraining the a parameters to be equal. The change in G^2 was not significant, $G^2(20) = 31.2$, ns. These analyses imply that the PCL–R items do not differ significantly in their ability to discriminate between levels of psychopathy in Scotland and North America. Thus, the PCL–R items have similar relevance to the diagnosis of psychopathy in Scotland as they have in North America. However, certain of the characteristics become apparent only at more extreme levels of the underlying trait in Scotland as compared with North America. These differences are discussed in relation to the final IRT model (see below).

Differences in scaling. The presence of DIF between the two settings indicated that measurement of Θ in Scotland and North America was not on a common metric, and, thus, estimates of the underlying trait were not directly comparable. Fortunately, it is not necessary to have all items with equivalent parameters across groups to derive a common metric: Items that are invariant across settings can be used as anchors to establish a common metric across groups (see Reise et al., 1993, for a detailed account of this method).

Measurement invariance was tested on an item-by-item basis; items were constrained to have all three parameters equal in the two settings, and a series of models was estimated and compared with a baseline model in which the slopes, but not the difficulty parameters, were constrained to be equal. Items were added in

order of total difference in b_i parameters, one item at a time, until a significant increase in G^2 was found. Three items, namely, "pathological lying," "early behavior problems," and "criminal versatility" could be fitted to have all parameters equal across the settings without producing significant increases in G^2, $G^2(6) = 9.2$, ns. However, constraining the parameters of other items to be equal resulted in significant increases in G^2. It was concluded that the three items were invariant across groups and could be used as anchor items. The parameters for the final fitted model, in which slopes were equal across settings for all items and thresholds were equal for the three anchor items, are displayed in Table 2. The model fits the data well, predicting the observed pattern of responses for each item within 1%.

The item-characteristics curves for "glibness/superficial charm," "callous/lack of empathy," and "promiscuous sexual behavior" are plotted to illustrate some of the differences between Scotland and North America (see Figure 3). "Callous/lack of empathy" is the most discriminating of the three items in both settings, but there is a substantial difference in threshold between North America and Scotland, with the Scottish participants having to have high levels of the trait before scoring positively on this item. The difference in threshold is most marked for "glibness/superficial charm." "Promiscuous sexual behavior" is the least discriminating of these items in both settings, with, unusually, the thresholds being lower in Scotland than in North America. Having derived a model using data from both settings, it was now valid to compare estimates of psychopathic personality disorder across the two settings; although the PCL–R total scores were not directly comparable, the disorder was now measured on the same underlying metric Θ.

Assessing diagnostic cutoffs. In making any cross-cultural comparison, it is necessary to ensure not only that the same construct is being measured in the different settings but also that it is being measured on the same metric (Hulin, 1987; Reise et al., 1993; van de Vijver & Leung, 1997). It should be noted that the equating of three items did not result in large changes to the other parameters. This indicates that although the PCL–R scores have different meanings in the two settings—as a consequence of the relations between the PCL–R scores and Θ being different across settings—the metrics of the latent trait in use in the two cultures were already almost identical.

An analysis of variance on the estimated level of trait for each participant showed significant differences, $F(4, 2308) = 25.8, p < .001$, in the mean level of the trait across samples, with the United States samples having the highest mean (White prisoners, $\bar{\Theta} = 0.51$; Black prisoners, $\bar{\Theta} = 0.54$), the Scottish sample having the lowest mean ($\bar{\Theta} = 0.00$), and the Canadian samples having intermediate values (hospital patients, $\bar{\Theta} = 0.02$; prisoners, $\bar{\Theta} = 0.26$). (Note that MULTILOG estimates group means by having set one group mean to zero and all standard deviations to 1.0.) This variation can be explained in part by variations in sample composition. For example, within North America, it is normal for the mean PCL–R scores to be lower among forensic patients as compared with prisoners (Hare, 1991).

To examine these relations, we estimated the regression of Θ on PCL–R scores across settings. A common regression line was fitted for the North American data ($\Theta = -2.27 + 0.110$PCL–R, $r = .98$) and one for the Scottish data ($\Theta = -1.68 + 0.111$PCL–R, $r = .98$). Hare (1991) recommended two diagnostic cutoff points for use with the PCL–R; scores of 30 or above are regarded as being diagnostic of psychopathy, and scores of 20–29 are indicative of moderate degrees of psychopathic personality disorder. Entering two cutoff values (i.e., 20 and 30) into the

Table 2

Item Parameters for North America (NA) and Scotland (SC); Equal Slopes and Anchors

PCL–R item label	a	b_1 NA	b_1 SC	b_2 NA	b_2 SC
1. Glibness/superficial charm	1.3	−0.2	1.3	1.7	3.7
2. Grandiose sense of self-worth	1.4	−0.4	0.7	1.3	1.9
3. Need for stimulation	1.5	−1.3	−0.6	0.2	1.2
4. Pathological lying	**1.4**	**−0.6**	**−0.6**	**1.1**	**1.1**
5. Conning/manipulative	1.5	−0.5	−0.1	1.1	1.6
6. Lack of remorse or guilt	1.6	−1.6	−0.2	0.0	0.7
7. Shallow affect	1.6	−0.9	0.0	0.7	1.4
8. Callous/lack of empathy	1.9	−1.1	0.0	0.6	1.2
9. Parasitic lifestyle	1.0	−1.3	−0.1	1.4	1.5
10. Poor behavioral controls	1.0	−1.2	−0.8	0.6	0.9
11. Promiscuous sexual behavior	0.7	−0.8	−1.4	0.8	0.3
12. Early behavior problems	**1.0**	**−0.2**	**−0.2**	**1.0**	**1.0**
13. Lack of long-term goals	1.2	−1.4	−0.4	0.5	0.3
14. Impulsivity	1.3	−2.0	−0.5	−0.1	0.7
15. Irresponsibility	1.4	−1.9	−0.5	0.0	1.4
16. Failure to accept responsibility	1.0	−1.4	−1.0	0.6	1.9
17. Short-term marital relationships	0.7	0.8	4.3	2.3	6.2
18. Juvenile delinquency	0.8	−0.7	−1.5	0.5	−0.4
19. Revocation of conditional release	0.7	−1.4	0.5	−0.1	3.2
20. Criminal versatility	**0.9**	**−0.4**	**−0.4**	**1.5**	**1.5**

Note. Anchor items shown in boldface; all nonanchor items have significantly different b_i parameters across settings. The a parameter is a measure of the discriminating power of the item; b_1 and b_2 are the points of infection for a 0 rating and a 2 rating, respectively. PCL–R = Psychopathy Checklist—Revised.
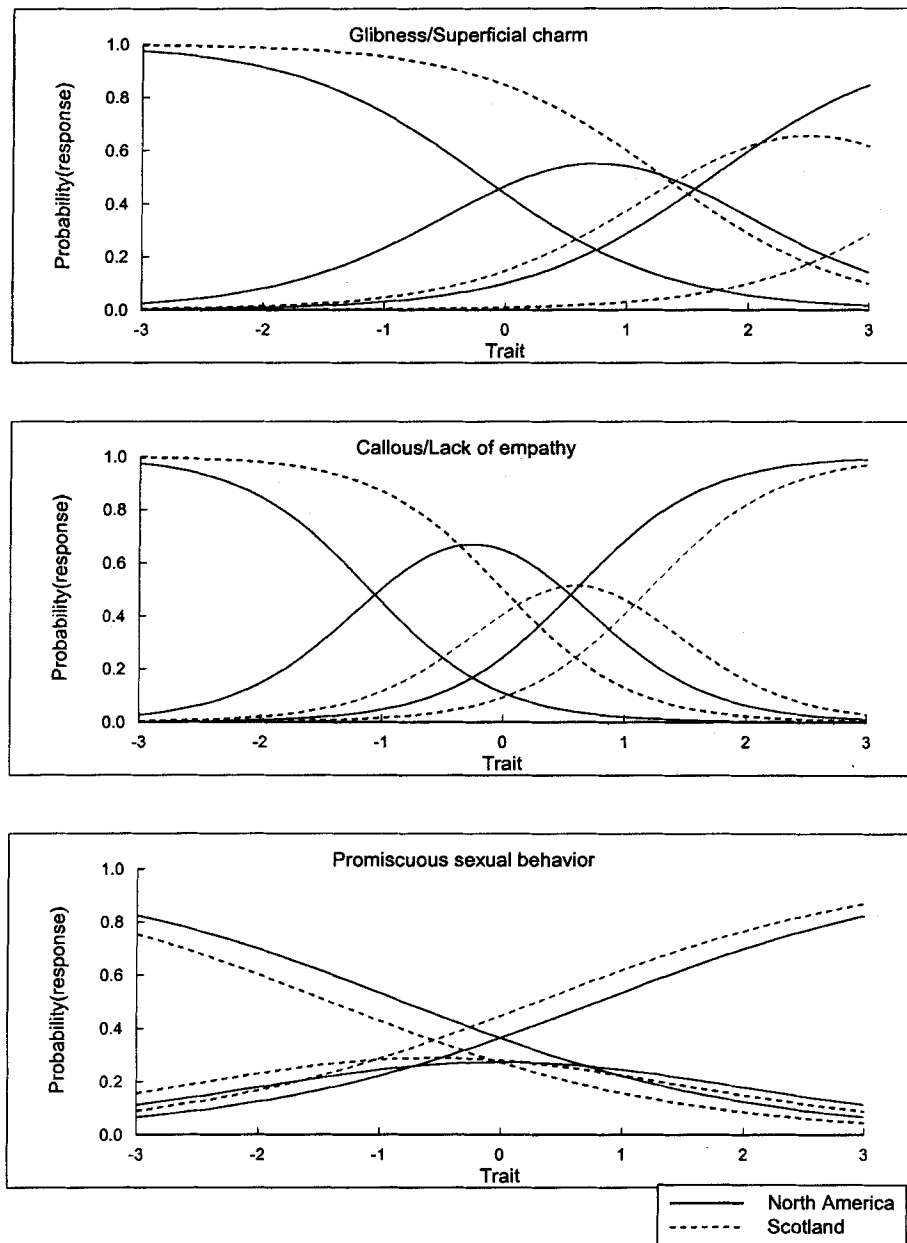
Figure 3. Item response curves derived from the North American and Scottish samples for three Psychopathy Checklist—Revised items, "glibness/superficial charm" (top), "callous/lack of empathy" (middle), and "promiscuous sexual behavior" (bottom).

regression equation for North America provides equivalent values of Θ of -0.05 and 1.06, respectively. Given that the anchoring procedure ensures that the estimated Θ is on the same metric in the two settings, by entering these values into the regression equation for Scotland it is possible to derive equivalent PCL–R score cutoffs; a cutoff of 30 in North America is equivalent to 25 in Scotland, and a cutoff of 20 is equivalent to 15.

Difference in prevalence. New prevalence estimates were calculated with the new diagnostic cutoffs of 15 and 25. Although the estimated prevalence in Scotland increases, the prevalence of

psychopathy in Scottish prison samples remains significantly lower than in North American samples (29% above 30); the prevalence in Scotland (8% above 25) being approximately one quarter of that observed in North American samples, $\chi^2(N = 2,309) = 70.9$, $p < .001$. This is a substantial difference.

Discussion

An initial exploratory study of this nature must raise as many questions as it answers. Nonetheless, these results provide some

empirical support for Hare's (1998) contention that the expression of psychopathy may be influenced by cultural processes.

In this discussion, we focus on five issues: the three specified in the preamble to the Results section and two broader issues, namely, how can cross-cultural research in this area be progressed and what are the practical implications of these findings?

The Relevance of PCL–R Items for Measuring Psychopathy in Scotland

The analyses confirm that the PCL–R is a good measure of psychopathic personality disorder in Scotland: All of the items contribute to the estimate of the trait, and there are different items that discriminate well, at different points, along the whole length of the trait. The results on the Scottish data set corroborate previous findings with North American data sets (Cooke & Michie, 1997; Cooke et al., in press). The fact that the slope parameters did not differ significantly across settings indicates that the disorder is defined by the same types of affective, behavioral, and interpersonal characteristics in Scotland as it is in North America. This implies that there is good cross-cultural generalizability of the construct and that the "pan-cultural core" of the disorder is essentially the same.

The most parsimonious IRT model suggests that differences in item response curves across the two settings are differences of extremity rather than differences of slope. Hulin (1987) indicated that nonequivalence of item curves underpinned by differences in slope is more profound than differences of extremity. Difference in the slope of an item across groups indicates that the item is less relevant to the underlying latent trait in one group as compared with the other group; this suggests that the item is culturally specific rather than culturally general. A difference in the extremity parameter (b_i) suggests that the item has similar relevance for the trait in both settings and that differences have to do with the scaling within items; the behavioral expression of the item has a greater or lesser range in one setting compared with the other.

Is There Evidence for the Suppression or Facilitation of Features of Psychopathy?

The findings that many of the features of the disorder do not become apparent in Scottish prisoners until high levels of the trait are achieved suggests that there may be culturally related processes that damp down, inhibit, or suppress the expression of these characteristics. As noted earlier, Weisz et al. (1987) argued that cultural characteristics may mold the expression of a disorder because some characteristics may be suppressed, whereas others may be accentuated.

The difference in b_i parameters between North America and Scotland varied across items; clearly, item differences will occur for a variety of reasons. Differences in items such as "revocation of conditional release" and "many short-term marital relationships" may reflect differences in the functioning of the criminal-justice system or differences in social practice, respectively (Fiske, 1995).

It could be speculated that the other large differences, for example, in items including "glibness/superficial charm" and "grandiosity," may be influenced by a social desirability response set. The acceptable range of the expression of a characteristic is likely to be influenced by cultural norms. It is our impression that the cultural norm in Scotland is against talking about one's abilities, prowess, or accomplishments—an exaggerated form of British reserve—unfortunately, no systematic comparison of cultural differences of this type is available in the literature comparing Scotland and North America. Differences of this type are known to exist when other cultures are compared (e.g., Barnlund, 1989; Lewin, 1948). Differences, if such exist, would be influential (Fiske, 1995).

Other cultural processes may be important. A core facet of psychopathy is the "selfish, callous, and remorseless use of others" (Hare, 1991): It is a facet underpinned by moral judgments. Fiske (1995) indicated that moral judgments are influenced to a significant degree by moral standards that are culturally transmitted through cognitive and affective processes.

There is an extensive body of evidence that supports the view that cultural factors influence behavioral, affective, and interpersonal characteristics. Indeed, the cross-cultural approach in psychology is predicated on the assumption that behavioral similarities within cultures and behavioral differences across cultures are developed and maintained through enculturation and socialization. A central explanatory construct in cross-cultural psychology is the "individualistic–collectivist" dimension (Berry, Poortinga, Segall, & Dasen, 1992). Individualistic cultures emphasize competitiveness and self-confidence; independence from others is encouraged, and temporary or short-lived relationships are common. By way of contrast, within collectivistic cultures, an individual's contribution and subservience to the social group is emphasized, the acceptance of authority is paramount, and continuous stable relationships are common. Within individualistic societies, cultural transmission is likely to enhance grandiosity, glibness and superficiality, promiscuity and multiple marital relationships, as well as a lack of responsibility for others. By way of contrast, cultural transmission within collectivistic societies will bear down on self-expression and promote stable family and group relationships. The competitiveness inherent in individualistic societies not only produces higher rates of criminal behavior but also leads to an increased use of Machiavellian behavior, in particular, an increase in the use of deceptive, manipulative, and parasitic behavior (Christie & Geis, 1970; Mealey, 1995; Wilson & Herrnstein, 1985).

There is empirical evidence from cognate constructs that supports this view. Compton et al. (1991) compared epidemiological data collected in community sites in Taiwan and in the United States of America and thereby compared prototypically collectivistic and individualistic societies. They observed substantial differences in the rates of APSD; the rates ranged from 0.10% to 0.22% in the Taiwanese sites, compared with a range of 1.49% to 5.66% in sites in the United States. These differences may represent real differences in the prevalence of the disorder, differences in the expression of the disorder, or a combination of both.

Data from the Epidemiological Catchment Area study suggest that within the United States, cultural pressures may influence the prevalence of ASPD. Robins et al. (1991) predicted that the lifetime prevalence of this disorder would increase from 3.7% to 6.4% by the time the members of the youngest cohort attained 30 years of age. Paris (1993) contended that changes of this nature can be attributed to a reduction in frequency of stable relationships, together with a weakening of the elements of the social fabric that suppress traits such as impulsivity.

Other evidence that suggests that psychopathy is influenced by the processes of enculturation and socialization may be found in the evidence for gender differences in the prevalence of psychopathy and related disorders. Within North America, differences in the prevalence of antisocial personality disorders as measured by the PCL–R and by *DSM–III, DSM–III–R,* and *Research Diagnostic Criteria* suggest that there are substantial gender differences in the prevalence of these disorders (Rutherford, Alterman, Cacciola, & Snider, 1995; Rutherford, Cacciola, Alterman, & McKay, 1996). Indeed, these authors' suggestion that the diagnostic cutoff should be lowered when rating women implies that they believe that there is a generalized suppression of PCL–R scores.

Why Is There a Comparatively Low Prevalence of Psychopathy in Scottish Prisons?

The results detailed above suggest that even after correcting for differences in the metric of the PCL–R, there still remains a substantially lower prevalence of psychopathy among Scottish prisoners. This finding must be couched with caveats about the equivalence of samples, although the magnitude of the difference (almost 4 to 1) is probably too great to be explained entirely by sampling variability. During the relevant period, the United States incarcerated between 5 to 8 times as many people per 100,000 than did Scotland (Council of Europe, 1995). If it is assumed that psychopaths are more likely to be imprisoned than are nonpsychopaths, then a country that imprisons a smaller proportion of its citizens should have proportionately more psychopaths in prison. Thus, the substantially lower prevalence in Scotland is contrary to that which is expected.

One intriguing possibility to explain the apparently lower prevalence of the disorder in Scotland—other than differences in cultural pressure toward psychopathic behavior—may be that psychopaths migrate out of Scotland. Mealey (1995) argued that psychopaths migrate. Why should this be the case? Wilson and Herrnstein (1985) argued that criminals, in general, are more likely to migrate to large modern cities. In such settings, the relationships among neighbors tend to be superficial. Such settings provide many targets for the predator; the predator is able to attack, extort, or steal from victims with little danger that the victim will recognize them. Clearly, such an environment would be highly fitting for the psychopath (Hare, 1993).

The need to seek out new opportunities and victims to exploit may go some way to explain psychopaths' tendency to migrate; however, other characteristics of the disorder may also play a significant role in their tendency to migrate. For example, impulsivity and the need for stimulation, together with their failure to form and maintain long-lasting relationships and their lack of realistic, long-term goals, may all serve to feed psychopaths' tendency to migrate (Cooke, 1998).

To test the hypothesis that Scottish psychopaths migrate, criminal records were obtained for crimes committed in England and Wales. Of the sample, 33% had a conviction in England and Wales. Those with one or more convictions in England and Wales had significantly higher scores on the total PCL–R score than did those without convictions in England and Wales, $t(223) = 3.28$, $p = .001$. This association cannot be explained merely in terms of psychopaths committing more crimes; the association persists when the sample is split at the median in terms of number of adult convictions, low group, $t(107) = 2.2$, $p = .04$; high group, $t(116) = 2.0$, $p = .03$.

Developing Cross-Cultural Research on Psychopathy

The PCL–R probably represents the most useful tool for pursuing issues of cross-cultural generalizability; within the North American context, and increasingly within the European context, the PCL–R has been shown to have impressive convergent and concurrent validity. The current study must be regarded as a first attempt to apply appropriate psychometric techniques to the question of the cross-cultural generalizability of the construct. What future steps should be taken to improve our understanding?

Perhaps the first step is to consider, in much more detail, the possibility that rater nationality has an impact on the ratings. An ongoing study is examining the quantitative effects of the raters' nationality and the effects of the prisoners' nationality on PCL–R scores.

One of the primary advantages of IRT methodology is that the parameter estimates are independent of the samples from which they are derived (Embretson, 1996). This is clearly not the case for prevalence estimates. The large difference in prevalence observed between Scotland and North America may, in part, be attributed to differences in sampling. One method for examining the impact of sampling on prevalence estimates would be to generate matched samples of Scottish and North American prisoners from existing data sets: Matching variables might include gender, age, types of offense, socioeconomic status, and educational level. Bontempo (1993) has argued for this approach in the cross-cultural context, indicating that researchers should initially consider variables that describe individuals and then move toward variables that describe the cultures or settings that are being compared. A multiple regression strategy can be adopted within which the individual level variables are introduced into the model initially, and, at the next stage, the culture variables are stepped into the model. In the ideal case, all differences should then be explained.

van de Vijver and Leung (1997) argued that the type of psychological difference study reported in this article is an appropriate approach to adopt in the early exploration of putative differences across cultures. However, if our understanding is to progress, it is necessary to explore the variables that may explain the apparent differences. Within the context of psychopathy, at least three areas should be explored, namely, level of self-disclosure, interpersonal style, and individualism versus collectivism.

All ratings on the PCL–R are to some extent dependent on what the participant tells the interviewer or has previously told others. To some extent, therefore, self-disclosure could influence PCL–R scores. There is evidence that indicates that PCL–R scores are dependent, to some degree, on the amount of information that is available to the rater. For example, Alterman, Cacciola, and Rutherford (1993) found that PCL–R scores obtained with interview alone were lower than those obtained when collateral information was added to the interview information.

It is known that levels of self-disclosure vary across cultures, with Americans disclosing more than Chinese (Chen, 1995), Germans (Lewin, 1948) and Japanese (Barnlund, 1989). It is plausible that North Americans disclose more than Scots, as well. Systematic study of self-disclosure across countries would perhaps cast a light on differences in the PCL–R ratings.

Psychopathic individuals are described as having an interpersonal style that is dominant, forceful, arrogant, and manipulative. Perhaps, there are differences between Scotland and North America that can be attributed to differences in interpersonal style, with suppression of some of these characteristics being affected by cultural pressure. This is an empirical question. Fortunately, we now have a methodology available for measuring the interpersonal style of psychopaths and others in correctional institutions—the Interpersonal Measure of Psychopathy (Kosson, Steurwald, Forth, & Kirkhart, 1997). Ongoing research is examining the relation between PCL–R scores and scores on this new measure in a Scottish prison sample to determine whether differences in interpersonal style emerged in this sample compared with North American samples.

Earlier, it was argued that differences in the level of individualism may affect the prevalence of psychopathic characteristics. Unfortunately, we have not been able to identify any studies that consider the position of Scotland on the individualism–collectivism dimension. Studies aggregate Scotland with the rest of the United Kingdom, and there are good reasons to suspect that there are differences between Scotland and the rest of the United Kingdom in respect to this dimension. The location of Scotland on this measure should be explored.

Practical Implications of the Putative Differences

Although cross-cultural research in general has considerable psychological and theoretical interest, cross-cultural research on the PCL–R has, in addition, considerable practical and ethical importance. Almost uniquely for a psychological instrument, the PCL–R can have a dramatic effect on the lives of those assessed by it. The PCL–R is an integral part of an increasing number of risk assessment protocols; indeed, the PCL–R is being used in assessments for release from prisons and forensic hospitals within Canada and the United States (Hare, 1998). There is a growing trend for the instrument to be adopted in a number of European countries. Given the saliency of this instrument in important decisions about individual liberty, it is imperative that more detailed cross-cultural analyses are carried out.

References

Alterman, A. I., Cacciola, J. S., & Rutherford, M. J. (1993). Reliability of the Revised Psychopathy Checklist in substance abuse patients. *Psychological Assessment, 5*, 442–448.

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

Barnlund, D. C. (1989). *Communicative styles of Japanese and Americans: Images and realities.* Belmont, CA: Wadsworth.

Barrett, P., & Eysenck, S. B. G. (1984). The assessment of personality factors across 25 countries. *Personality and Individual Differences, 5*, 615–632.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications.* Cambridge, England: Cambridge University Press.

Bontempo, R. (1993). Translation fidelity of psychological scales: An item response theory analysis of an individualism–collectivism scale. *Journal of Cross-Cultural Psychology, 24*, 149–166.

Chen, G. (1995). Differences in self-disclosure patterns among Americans versus Chinese: A comparative study. *Journal of Cross-Cultural Psychology, 26*, 84–91.

Christie, R., & Geis, F. L. (1970). *Studies of Machiavellianism.* London: Academic Press.

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

Compton, W. M., Helzer, J. E., Hwu, H. G., Yeh, E. K., McEvoy, L., Tipp, J. E., & Spitznagel, E. L. (1991). New methods in cross-cultural psychiatry: Psychiatric illness in Taiwan and the United States. *American Journal of Psychiatry, 148*, 1697–1704.

Conoley, J. C., & Impara, J. C. (1995). *12th mental measurement yearbook.* Lincoln, Nebraska: Buros Institute.

Cooke, D. J. (1989). Containing violent prisoners: An analysis of the Barlinnie Special Unit. *British Journal of Criminology, 29*, 129–143.

Cooke, D. J. (1994). *Psychological disturbance in the Scottish Prison System: Prevalence, precipitants and policy.* Edinburgh, Scotland: Scottish Prison Service.

Cooke, D. J. (1995a). Psychological disturbance in the Scottish prison system: A preliminary account. In G. Davie, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (pp. 495–509). Berlin, Germany: Walter de Grutyer.

Cooke, D. J. (1995b). Psychopathic disturbance and offending in the Scottish prison population. *Psychology, Crime and Law, 2*, 101–118.

Cooke, D. J. (1996). Psychopathic personality in different cultures. What do we know? What do we need to find out? *Journal of Personality Disorders, 10*, 23–40.

Cooke, D. J. (1998). Cross-cultural aspects of psychopathy. In D. J. Cooke, A. E. Forth, & R. D. Hare. (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 13–45). Dordrecht, the Netherlands: Kluwer Academic.

Cooke, D. J., & Michie, C. (1997). An item response theory evaluation of Hare's Psychopathy Checklist—Revised. *Psychological Assessment, 9*, 2–13.

Cooke, D. J., Michie, C., Hart, S., & Hare, R. D. (in press). The functioning of the clinical version of the Psychopathy Checklist: An item response theory analysis, *Psychological Assessment.*

Council of Europe. (1995). *Penological information bulletin,* Nos. 19–20.

Dinges, N. G., Atlis, M. M., & Vincent, G. M. (1997). Cross-cultural perspectives on antisocial behavior. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of antisocial behavior* (pp. 463–473). New York: Wiley.

Draguns, J. G. (1973). Comparison of psychopathology across cultures: Issues, findings, directions. *Journal of Cross-Cultural Psychology, 4*, 9–47.

Draguns, J. G. (1986). Culture and psychopathology: What is known about their relationship? *Australian Journal of Psychology, 38*, 329–338.

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

Fiske, A. P. (1995). The cultural dimensions of psychological research: Method effects imply cultural mediation. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research method and theory: A festschrift honoring Donald W. Fiske* (pp. 271–294). Hillsdale, NJ: Erlbaum.

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

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

Hare, R. D. (1998). The Hare PCL–R: Some issues concerning its use and misuse. *Legal and Criminological Psychology, 3*, 101–119.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning.* Hillsdale, NJ: Erlbaum.

Hulin, C. L. (1987). A psychometric theory of evaluations of item and scale translation: Fidelity across languages. *Journal of Cross-Cultural Psychology, 18*, 115–142.

Kosson, D. S., Steurwald, B. L., Forth, A. E., & Kirkhart, K. J. (1997). A new method for assessing the interpersonal behavior of psychopaths: Preliminary validation studies. *Psychological Assessment, 9*, 89–101.

Lewin, K. (1948). Some socio-psychological differences between the United States and Germany. In G. Lewin (Ed.), *Resolving social conflicts: Selected papers on group dynamics* (pp. 1935–1946). New York: Harper.

Lewis-Fernandez, R., & Kleinman, A. (1994). Culture, personality, and psychopathology. *Journal of Abnormal Psychology, 103,* 67–71.

Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences, 18,* 523–599.

Mellenbergh, G. J. (1996). Measurement precision in test score and item response models. *Psychological Methods, 1,* 293–299.

Murphy, J. M. (1976, March 12). Psychiatric labeling in cross-cultural perspective: Similar kinds of disturbed behavior appear to be labeled abnormal in diverse cultures. *Science, 191,* 1019–1028.

Paris, J. (1993). Personality disorders: A biopsychosocial model. *Journal of Personality Disorders, 7,* 255–264.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114,* 552–566.

Robins, L. N., Tipp, J., & Przybeck, T. (1991). Psychiatric disorders in America. In L. N. Robins & D. A. Regier (Eds.), *Antisocial personality disorder* (pp. 258–290). New York: Free Press.

Rutherford, M. J., Alterman, A. I., Cacciola, J. S., & Snider, E. C. (1995). Gender differences in diagnosing antisocial personality disorder in methadone patients. *American Journal of Psychiatry, 152,* 1309–1316.

Rutherford, M. J., Cacciola, J. S., Alterman, A. I., & McKay, J. R. (1996). Reliability and validity of the revised Psychopathy Checklist in women methadone patients. *Assessment, 3,* 145–156.

Sanchez, J. (1986). Social crises and psychopathy: Towards a sociology of psychopathy. In W. H. Reid, D. Dorr, J. I. Walker, & J. W. Bonner (Eds.), *Unmasking the psychopath: Antisocial personality disorders and related syndromes* (pp. 78–97). New York: Norton.

Steinberg, L., & Thissen, D. (1996). Uses of item response theory and the testlet concept in the measurement of psychopathology. *Psychological Measurement, 1,* 81–97.

Thissen, D. (1991). *MULTILOG user's guide (version 6)* [Computer software]. Mooresville, IN: Scientific Software.

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

Weisz, J. R., Suwanlert, S., Chaiyasit, W., & Walter, B. R. (1987). Over- and undercontrolled referral problems among children and adolescents from Thailand and the United States: The Wat and Wai of cultural differences. *Journal of Consulting and Clinical Psychology, 55,* 719–726.

Wilson, J. Q., & Herrnstein, R. J. (1985). *Crime and human nature.* New York: Simon & Schuster.