Complementary approaches to the assessment of personality disorder

The Personality Assessment Schedule and Adult Personality Functioning Assessment compared

JONATHAN HILL, HAZEL FUDGE, RICHARD HARRINGTON, ANDREW PICKLES and MICHAEL RUTTER

Background Current concepts and measures of personality disorder are in many respects unsatisfactory.

Aims To establish agreement between two contrasting measures of personality disorder, and to compare subject–informant agreement on each. To examine the extent to which trait abnormality can be separated from interpersonal and social role dysfunction.

Method Fifty-six subjects and their closest informants were interviewed and rated independently. Personality functioning was assessed using a modified Personality Assessment Schedule (M–PAS), and the Adult Personality Functioning Assessment (APFA).

Results Subject–informant agreement on the M–PAS was moderately good, and agreement between the M–PAS and the APFA, across and within subjects and informants, was comparable to that for the M–PAS. This was equally the case when M–PAS trait plus impairment scores and trait abnormality scores were used.

Conclusions The M–PAS and the APFA are probably assessing similar constructs. Trait abnormalities occur predominantly in an interpersonal context and could be assessed within that context.

The concepts and measurements of personality disorder are, in many respects, problematical (Rutter, 1987; Hill & Rutter, 1994; Hill et al, 1995). Difficulties with measurement have included lack of agreement among measures, low subject–informant agreement, substantial comorbidity and lack of distinctiveness of disorders (Zimmerman et al, 1988; Zanarini et al, 1990; Oldham et al, 1992; Riso et al, 1994; Dolan et al, 1995; Tyrer & Johnson, 1996). Clinicians do not generally use existing typologies of personality disorder in assessing patients (Westen, 1997). The Adult Personality Functioning Assessment (APFA; Hill et al, 1989) was developed to assess personality functioning as a pattern of interpersonal and social role performance over substantial periods of time. It appears to assess an underlying latent variable of personality dysfunction, and has good interrater reliability and subject–informant agreement (Hill et al, 1989, 1995). This study was designed to explore similarities and differences with the modified Personality Assessment Schedule (M–PAS; Tyrer & Alexander, 1979).

The PAS

The PAS has been the most extensively used British standardised assessment of personality disorders. It has been shown to possess adequate interrater and test–retest reliability (Tyrer & Alexander, 1979; Tyrer et al, 1983) and to be predictive of treatment outcome (Tyrer & Seivewright, 1988). It was, therefore, appropriate for a comparison with the APFA. The PAS is a standardised interview in which the interviewee is asked for information related to 24 personality characteristics, and where the answer is positive he or she is asked to provide examples of relevant behaviours. Ratings for each trait are made on a nine-point scale from 0–8, on which the ratings of 0–3 are trait accentuations in the absence of impairment of social functioning or distress to the subject or those around him or her. The scores are combined using formulae described by Tyrer et al (1988) to derive either 13 personality disorder categories, or four summary categories.

The M–PAS

Some modifications were made to the PAS by the authors so that additional questions could be addressed while retaining the scoring method of the original instrument. The main purposes of the alterations were: (a) to provide questions that might improve subject–informant agreement in the reporting of traits; (b) to enable trait abnormality and impairment to be assessed separately; and (c) to add traits that might be relevant in studies of the families of autistic individuals (Piven et al, 1994). This modified instrument will be referred to as the M–PAS throughout the paper. The term 'trait' will be reserved for persistent cognitive/emotional/behavioural patterns assessed separately from impairment in the M–PAS.

Low subject–informant agreement has been a major problem for personality disorder measures (Zimmerman et al, 1988; Riso et al, 1994) and it is likely that there are several contributing factors (Hill et al, 1995). One possibility is that subjects and informants may make different attributions of the same behaviours. For instance, behaviour reflecting suspiciousness might be recognised by both subject and informant. However, the subject may see it as the reasonable response to a perceived threat while to the informant it is an overreaction to a trivial event. This might be overcome where the subject interview asks about his or her experience of other people, and the informant interview about the subject's perception of others. Thus, in the M–PAS for subjects the 'suspiciousness' question was changed from "In general how well do you get on with other people? Do you normally trust them or are you suspicious of them, at least at first?" to "In general do you find that people are to be trusted?" The follow-up question then enquired about the basis of the subject's experiences, leaving the investigator to form the judgement as to whether this amounted to significant level of inappropriate suspiciousness. Similar changes, reflecting the likelihood that subjects and informants might interpret the same behaviours differently, were made to most of the PAS questions.

The relationship between trait and impairment is fundamental. In the PAS it is assumed that "each person possesses a small and distinct group of primary traits that persist over time and exhibit a high degree of consistency across situations" (Millon, 1987) and that those with personality disorder ". . . would develop personal and social dysfunction as a consequence of excessive prominence of these traits . . ." (Tyrer & Alexander, 1988). Thus, trait and impairment are conceptualised as separable and by implication could be measured independently. Furthermore, there might be variability in the extent of association between abnormal trait and impairment and investigation of factors associated with such variation could be important. For these reasons the ratings of trait and impairment were separated in the M–PAS. In order to do this four dimensions of trait abnormality were identified: (a) threshold; (b) intensity; (c) persistence and (d) pervasiveness. A rating of 'zero' was made where the trait was absent and of 'one' where it was present and either only affected feelings or was associated with behaviours that were normal and appropriate. The rating of 'two' required that behaviour was affected and either that there was a low threshold for the response or that there was a high intensity of behaviour. Where there was both low threshold and high intensity a rating of at least 'three' was made. Ratings from 'three' to 'six' were all at the same level of threshold and intensity but reflected whether the behaviours were seen continually or episodically, and whether they were pervasive. The numbers of subjects scoring on each of the four severe trait ratings was low and so for analyses presented in this paper ratings in the range 'three' to 'six' were collapsed yielding a trait scale of 'zero' to 'three'. Impairment was rated on a 'zero' to 'five' scale where 'one' reflected 'some' impairment and points between 'two' and 'five' reflected significant impairment of different severity and pervasiveness. For the analyses presented here 'zero' and 'one' were coded as 'zero' so that the ratings from 'one' to 'four' represented varying levels of significant impairment.

These two scales of trait (0–3) and impairment (0–4) can be handled separately or combined to create a 0–7 scale. This closely parallels the points on Tyrer's scale, omitting the score of 'eight' which could be rated on the PAS only where impairment was so severe that the person was incapable of independent functioning.

The APFA

The APFA has been described in detail in previous papers (Hill *et al*, 1989, 1995). In brief, it provides a standardised assessment of a person's functioning in a range of social domains, with the aim of both identifying dysfunction that is specific to particular domains and measuring social dysfunction that is pervasive. Ratings of 'zero' to 'five' are made in each of six domains and the sum of these scores is taken to reflect the severity and pervasiveness of dysfunction. Detailed rating rules, a dictionary of examples and training, ensure that the individual's contribution to functioning, over substantial periods of time, and where possible free of DSM–IV Axis I symptoms, is rated.

This study had three aims. The first was to establish subject and informant agreement using the modified PAS, and to compare this with that obtained with the APFA (Hill *et al*, 1995). The second was to examine agreement between the M–PAS and the APFA, in order to establish whether these two contrasting measures appear to assess similar constructs. Third, the differences between the two measures were exploited to examine the extent to which trait abnormality can be separated from interpersonal and social role dysfunction.

METHOD

Sample

The sample was the same as that used in an examination of subject and informant agreement using the APFA (Hill *et al*, 1995). In brief, it was designed to provide an adequate spread of functioning over which to examine the performance of the PAS and the APFA, although all of the subjects were able to support themselves in the community. There were 42 parents (21 men, 21 women) of children attending the Maudsley Hospital Children's Department, and 14 patients (six men, eight women) attending the Adult Out-Patient Department of the Maudsley Hospital with long-standing, non-psychotic problems. Thus, there were in all 56 subjects (27 men, 29 women) with an age range of 23–69 years, a mean age of 41 years, and a median age of 40 years.

Subject–informant comparison

In this paper subject–informant comparisons using the M–PAS are presented, as

are comparisons of the APFA and the M–PAS between subjects and informants, thus providing a stringent test of agreement between the instruments. Details of the establishment of the subject–informant pairs were given in a previous paper (Hill *et al*, 1995). In brief, where the subject was married or cohabiting the spouse or cohabitee acted as the informant, and all married or cohabiting subjects acted both as subject and informant. Where the subject was living alone the identification of the best available informant was done through discussion with the subject after the purpose of the study had been explained. This study was designed so that blindness between subject and informant interviews and ratings was ensured. Thus, no one interviewer interviewed both subject and informant. As a result of the design constraints arising from the comparison of two forms of the SADS–L (Harrington *et al*, 1988) within one source of information, the interviewers were not balanced over subject and informant sources. In all, 19 subject and 10 informant interviews were carried out by R.H., 19 subject and 10 informant interviews by J.H. and 18 subject and 36 informant interviews by H.F. The possibility that this lack of balance was leading to interviewer effect being confounded with subject–informant effects was excluded by checking that the pattern of findings was consistent across interviewer pairs.

RESULTS

Interrater reliability

Interrater reliabilities for the trait plus impairment scores were assessed for three raters (R.H., J.H., H.F.) who each rated audio tapes of 21 interviews. Intra-class correlation coefficients (ICCs) are a widely used measure of agreement (Bartko & Carpenter, 1976) and these are shown in Table 1 together with 95% confidence intervals (CIs).

Subject–informant agreement

Subject–informant agreement was estimated for the combined trait and impairment scores and for traits only, using the M–PAS. When 0–7 scales were used, calculated from the sum of the traits and impairments, as described earlier, the intra-class correlations and 95% CIs shown in Table 1 were obtained. The majority of the ICCs were low. For the personality

attributes 'lability' and 'aloofness' this was probably attributable to low interrater reliability, but for the remainder there appeared to be low subject–informant agreement. A very similar pattern was seen when the traits scored on a 0–3 scale were entered, and only the ICCs for aloofness, impulsiveness, aggressiveness, irresponsibility, childishness, resourcelessness and dependence were significant at $P < 0.01$.

The PAS yields scores and categories for 13 types of personality disorder, and these can be combined to form four main disorders. The scores were calculated using the combined trait and impairment scales according to the method described by Tyrer *et al* (1988) and subject–informant agreement computed. ICCs for each of the 13 disorders, and for the four main disorders are shown in Table 2, as are the Kappa statistics for these disorders where there were either subject or informant based ratings of disorder in over 10% of cases. All of the disorders that contribute to the overall 'anti-social disorder' showed moderate and significant levels of subject–informant agreement (sociopathic, explosive and sensitive–aggressive) and this was the case in two out of the three that contribute to the dependent disorder (passive–dependent and histrionic). The sum of all of the M–PAS trait plus impairment scores may be taken as an overall index of extensiveness and severity of personality dysfunction, and the ICC comparing subject and informant total scores was 0.40 (95% CI 0.16–0.59, $P < 0.001$).

A category of personality disorder was required using trait scores only, in order to make comparisons with personality disorder defined by M–PAS traits plus impairment, and with the APFA. Trait ratings of 'three' were used as these had been defined as reflecting high severity and low threshold, and hence were likely to be most appropriate to the identification of personality disorder. Examination of the distribution of the ratings indicated that if a cut-off of four or more trait ratings of 'three' were used, this would yield 13 individuals with personality disorder based on subject interviews and 14 from informant interviews. This was similar to the number that was generated using Tyrer's rules for deriving personality disorder on the basis of trait plus impairment scores. To what extent were they identifying similar individuals? On the basis of subject ratings there were five disagreements between the two methods with a Kappa of 0.73 ($P < 0.001$,

Table 1 Estimates of interrater reliability and subject–informant agreement for M–PAS trait plus impairment scores

Personality attribute	Interrater reliability ($n=21$) ICC (95% CI)		Subject–informant agreement ($n=56$) ICC (95% CI)	
Pessimism	0.63	(0.30–0.83)	0.19	
Worthlessness	0.52	(0.14–0.77)	0.00	
Lability	0.32		0.29	
Anxiousness	0.75	(0.49–0.89)	0.07	
Suspiciousness	0.72	(0.44–0.88)	0.16	
Introspection	0.69	(0.39–0.86)	0.01	
Shyness	0.61	(0.30–0.83)	0.11	
Aloofness	0.31		0.21	
Sensitivity	0.75	(0.49–0.89)	0.07	
Vulnerability	0.40	(0.00–0.70)	−0.04	
Irritability	0.71	(0.42–0.87)	0.27	
Impulsiveness	0.76	(0.49–0.89)	0.44	(0.21–0.63)
Aggressiveness	0.81	(0.60–0.92)	0.56	(0.35–0.71)
Callousness	0.88	(0.74–0.95)	0.33	(0.08–0.60)
Irresponsibility	0.82	(0.61–0.92)	0.48	(0.25–0.65)
Childishness	0.32	(−0.12–0.65)	0.37	(0.13–0.57)
Resourcelessness	–		0.26	(0.01–0.49)
Dependence	–		0.57	(0.37–0.72)
Submissiveness	0.60	(0.25–0.81)	0.00	
Conscientiousness	0.50	(0.22–0.80)	−0.08	
Rigidity	0.70	(0.41–0.87)	−0.02	

Intraclass correlation coefficients (ICCs) are given where the characteristic was rated in 10% of subjects by at least one rater. 95% CIs are given for ICCs that were significant at $P < 0.01$.

Table 2 Subject–informant agreement for M–PAS personality disorder categories (Kappas) and total scores (intraclass correlation coefficients (ICCs))

	Kappa	ICC (95% CI)	
13 personality disorder categories			
Sociopathic	0.39**	0.57***	(0.37–0.72)
Passive–dependent		0.54***	(0.33–0.70)
Anankastic		−0.05	
Schizoid		0.08	
Explosive	0.51***	0.58***	(0.38–0.73)
Sensitive–aggressive	0.29*	0.48***	(0.25–0.65)
Histrionic	0.30*	0.52***	(0.31–0.69)
Asthenic		0.12	
Anxious		0.15	
Paranoid		0.09	
Hypochondriacal		0.41***	(0.18–0.61)
Dysthymic		0.18	
Avoidant		0.11	
Four groups			
Antisocial	0.56***	0.56***	(0.35–0.71)
Dependent	0.20	0.43***	(0.20–0.62)
Inhibited		0.19	
Withdrawn	0.24*	0.06	

The significance of the Kappa and ICCs are indicated by *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. Kappas were calculated where personality disorder was rated either for subject or informant in at least 6/56 cases.

s.e.=0.112), and for informant-based ratings, 10 disagreements and a Kappa of 0.44 ($P<0.001$, s.e.=0.147).

It has been argued (Tyrer, 1987) that subjects with personality difficulties may be unable to describe their own deviant traits and so are likely to under-report when compared with informant accounts of their functioning. Possible under-reporting of deviant personality characteristics by subjects was examined by comparing the means of subjects and informants item by item. In comparisons of the 24 personality attributes subject interviews led to significantly higher ratings on suspiciousness (two-tailed t-test, $P<0.05$) and submissiveness (two-tailed t-test, $P<0.01$). In a comparison of the scores for the 13 personality disorder types the subject-based schizoid mean score was significantly higher than that of the informant-based score (two-tailed t-test, $P<0.05$) and there were no significant differences in the mean scores for the four principal personality disorder types. Given that two out of the three 'significant' differences were in the opposite to the predicted direction and that around two differences at $P<0.05$ could have been expected by chance there was no evidence that subjects underestimated their deviance.

Relationship between M–PAS and APFA

While the M–PAS and APFA are in many respects designed to perform different tasks, both may be used to identify presence or absence of personality disorder and their total scores may be taken to reflect severity of disorder in a similar fashion. Comparison of M–PAS and APFA ratings derived from the same source (subject or informant) is limited because the interviews were not carried out blind, to each other and,

therefore, are open to an over estimate of agreement. By contrast, comparisons of subject APFA scores with informant M–PAS scores and subject M–PAS scores with informant APFA scores are particularly severe tests in which subject–informant differences are likely to place a ceiling on the agreement that can be achieved between the instruments. Agreement between the PAS and APFA within and between subjects and informants is shown in Table 3.

The top row shows the agreement between the total M–PAS trait plus impairment scores and total APFA level scores. The agreement was moderately good, even when the different instruments were compared across subject and informant – the correlations of 0.48 and 0.59 are comparable to, if not better than, the correlation coefficient of 0.40 that was obtained for the subject–informant agreement using the M–PAS. It is evident in the second row that agreement between the M–PAS and APFA was very similar when the sum of the M–PAS traits (without impairment) was used. Agreement for presence or absence of personality disorder is shown in the third and fourth rows of the table. The Kappas across subject and informant were modest but similar to that of 0.43 for agreement using the M–PAS.

The figures in brackets in Table 3 refer to agreement between instruments after the 13 subjects with APFA ratings accompanied by symptoms had been removed. All but one of the Kappas showed modest improvements which were due predominately to a reduction in the number of disagreements arising from APFA scores above the cut off and absence of personality disorder as assessed on the M–PAS. Thus, within the limitations of the small numbers there was no evidence that the M–PAS was vulnerable to the effects of psychiatric symptoms.

DISCUSSION

Subject–informant agreement on the M–PAS

In this examination of inter-rater reliability of the M–PAS we found similar levels to those reported elsewhere. Subject–informant agreement on M–PAS totals of traits plus impairment scores and on presence or absence of personality disorder was modest but better than that reported for DSM-based instruments (Oldham *et al*, 1992; Riso *et al*, 1994). However, in this study the M–PAS was administered after the APFA which has been shown to have good subject–informant agreement (Hill *et al*, 1995), and so these findings may not apply when the M–PAS is used independently. Apart from antisocial personality disorder, agreement on the type of personality disorder was low and this is in line with the findings of Zimmerman *et al* (1988) and Riso *et al* (1994). Why was the agreement so poor? Much of the lack of agreement was attributable to poor agreement on items on the M–PAS. At least five reasons for this may be proposed. First, the items may not be adequately operationalised so that subject and informant are not referring to the same behaviours. Second, there may be differences in the type of information that subject and informant make use of in giving accounts of traits. For instance when asked about anxiousness a subject is likely to base the reply on feelings, while the informant may be more likely to infer anxiousness from behaviour. Third, subject and informant may be sampling from different situations. This is likely to be the case where the subject refers to behaviours in intimate relationships to which the informant does not have access. Fourth, subject and informant may summarise information differently so that for instance one refers to most severe behaviours

Table 3 Agreement between M–PAS and APFA based estimates of personality dysfunction for scores and categories

Comparison	Subject (PAS and APFA)	Subject (PAS) informant (APFA)	Subject (APFA) informant (PAS)	Informant (PAS and APFA)
M–PAS variable type				
Sum of scores M–PAS trait plus impairment – correlations	0.60** (0.60**)	0.48** (0.51**)	0.59** (0.65**)	0.60** (0.67**)
Sum of scores M–PAS traits – correlations	0.60** (0.61**)	0.44** (0.51**)	0.54** (0.56**)	0.54** (0.51**)
Personality disorder M–PAS trait plus impairment – Kappas	0.51*** (0.58***)	0.26* (0.38*)	0.44*** (0.48***)	0.51*** (0.59***)
Personality disorder M–PAS trait – Kappas	0.55*** (0.61***)	0.28* (0.34*)	0.34** (0.26*)	0.555*** (0.57***)

The significance of the Kappa and intraclass correlation coefficients are indicated by *$P<0.05$, **$P<0.01$ and ***$P<0.001$. Figures in parentheses refer to agreement after removal of APFA ratings, 'accompanied by symptoms'.

and the other to typical behaviours. Finally, there may be a combination of effects of subject and informant psychopathology. As we showed in the study of APFA subject–informant agreement, subjects with high scores reflecting more personality dysfunction are less likely to have close informants and those available may also be deviant and hence more unreliable. This would be consistent with the increase of disagreement with increase in PAS scores.

Agreement between the M–PAS and the APFA

The agreement between M–PAS and APFA was as strong when comparing subject and informant as the agreement between subject and informant using the M–PAS. This may in part reflect the fact that the APFA was administered to the subject or informant before the M–PAS so may have influenced it, however, it suggests that the two instruments are, broadly speaking, assessing similar constructs. If that is the case it would support our previous proposal (Hill *et al*, 1989; Hill & Rutter, 1994; Hill *et al*, 1995) that persistent dysfunctional patterns of social role and interpersonal performance may be common to many of the personality disorder categories. It is also consistent with Westen's (1997) finding that clinicians identify difficulties with intimacy and with work as common personality problems. The implications for measurement are that when the APFA is used as a measure of overall functioning, either as a continuous or categorical variable it identifies personality disorder, while still enabling those questions relating to severity, pervasiveness and persistence of dysfunction that we have identified in previous papers to be addressed.

Relationship between abnormal traits and impairment

In the version of the PAS used in this study, the M–PAS, traits and associated impairment were rated separately, and subject–informant agreement was similar when traits alone and the sum of traits and impairment were compared. The agreement between M–PAS trait scores, rated without reference to impairment, and APFA scores was particularly striking. It could be that as interviewers had already assessed patterns of social role and interpersonal functioning in the APFA when they administered the PAS they were inadvertently including impairment in their rating of traits. However, the

additional rules had made it clear how each item was to be judged, and they did not refer to impairment. An alternative explanation could be that the distinction between trait and impairment is in many respects artificial. The majority of the traits refer to behaviours that are seen in an interpersonal context so that the description of the trait has to be abstracted from interpersonal examples. In the APFA by contrast the assessment may be seen as being of traits *in situ* where they are rated directly. For instance, if the subject or informant is asked about the trait of aggressiveness, he or she will sample or summate from aggressive interpersonal events in order to provide examples, while in the APFA patterns of relationships and social role functioning characterised by aggression will be rated.

Future studies

Our findings need to be set against the background of issues to be tackled in the measurement of personality disorder outlined at the beginning of the paper, and previously (Hill *et al*, 1995). It has been important to

show that agreement between two measures that take different approaches to the assessment of personality functioning is good. The differences can then be exploited in order to address key questions in the conceptualisation and measurement of personality disorder. Further studies will take the examination further by making systematic links with ICD- and DSM-based instruments such as the International Personality Disorder Examination (Loranger *et al*, 1987).

REFERENCES

Bartko, J. J. & Carpenter, W. T. (1976) On the methods and theory of reliability. *Journal of Nervous and Mental Disease*, **163**, 307–317.

Dolan, B., Evans, C. & Norton, K. (1995) Multiple Axis-II diagnoses of personality disorder. *British Journal of Psychiatry*, **166**, 107–112.

Harrington, R., Hill, J., Rutter, M., et al (1988) The assessment of lifetime psychopathology: a comparison of two interviewing styles. *Psychological Medicine*, **18**, 487–493.

Hill, J., Harrington, R., Fudge, H., et al (1989) Adult personality functioning assessment (APFA): an investigator-based standardised interview. *British Journal of Psychiatry*, **155**, 24–35.

CLINICAL IMPLICATIONS

■ Personality disorder may be assessed in terms of interpersonal and social role performance.

■ Abnormal traits are generally seen within an interpersonal context.

■ Therapeutic approaches designed to improve interpersonal functioning may be helpful in the treatment of personality disorder.

LIMITATIONS

■ The sample included relatively few subjects with personality disorder.

■ The findings may not generalise to clinical samples with more psychopathology.

■ The findings may not generalise to those types of personality disorder that were not represented in the sample.

JONATHAN HILL, MRCPsych, Department of Psychiatry, Liverpool University; HAZEL FUDGE, BA, MRC Child Psychiatry Unit, Institute of Psychiatry, London; RICHARD HARRINGTON, FRCPsych, Department of Psychiatry, Manchester University; ANDREW PICKLES, PhD, MRC Child Psychiatry Unit, Institute of Psychiatry, London; MICHAEL RUTTER, FRCPsych, Social, Genetic and Developmental Research Centre, Institute of Psychiatry, London

Correspondence: Jonathan Hill, University Child Mental Health, Alder Hey Hospital, Eaton Road, Liverpool LI2 2AP

___ & Rutter, M. (1994) Personality Disorder. In *Child and Adolescent Psychiatry: Modern Approaches* (eds M. Rutter, E. Taylor & L. Hersov), pp. 688–696. Oxford: Blackwell.

___ , Fudge, H., Harrington, R., et al (1995) The Adult Personality Functioning Assessment (APFA): factors influencing agreement between subject and informant. *Psychological Medicine*, **25**, 263–275.

Loranger, A. W., Susman, V. L., Oldham, J. M., et al (1987) *International Personality Disorder Examination (PDE): A Structured Interview for DSM–III–R and ICD–10 Personality Disorders, WHO/ADAMHA Version.* White Plains, NY: New York Hospital–Cornell Medial Center, Westchester Division.

Millon, T. (1987) On the nature of taxonomy in psychopathology. In *Issues in Diagnostic Research* (eds C. G. Last & M. Hersen), pp. 3–85. New York: Plenum Press.

Oldham, J. M., Skodol, A. F., Kellman, H. D., et al (1992) Diagnoses of DSM–III–R personality disorders by two structured interviews: patterns of comorbidity. *American Journal of Psychiatry*, **149**, 213–220.

Piven, J., Wrozek, M., Landa, R., et al (1994) Personality characteristics of the parents of autistic individuals. *Psychological Medicine*, **24**, 783–795.

Riso, L. P., Klein, D. N., Anderson, R. L., et al (1994) Concordance between patients and informants on the Personality Disorder Examination. *American Journal of Psychiatry*, **151**, 568–573.

Rutter, M. (1987) Temperament, personality and personality disorder. *British Journal of Psychiatry*, **150**, 443–458.

Tyrer, P. (1987) Problems in the classification of personality disorder. *Psychological Medicine*, **17**, 15–20.

___ & Alexander, J. (1979) Classification of personality disorder. *British Journal of Psychiatry*, **135**, 163–167.

___ , Strauss, J. & Cicchetti, D. (1983) Temporal reliability of personality in psychiatric patients. *Psychological Medicine*, **13**, 393–398.

___ & Alexander, J. (1988) Personality Assessment Schedule. In *Personality Disorders* (ed. P. Tyrer), pp. 43–62. London: Wright.

___ , ___ & Ferguson, B. (1988) Appendix: Personality Assessment Schedule (PAS). In *Personality Disorders* (ed. P. Tyrer), pp. 140–167. London: Wright.

___ & Seivewright, H. (1988) Studies of outcome. In *Personality Disorders* (ed P. Tyrer), pp. 119–137. London: Wright.

___ & Johnson, T. (1996) Establishing the severity of personality disorder. *American Journal of Psychiatry*, **153**, 1593–1597.

Westen, D. (1997) Divergences between clinical and research methods for assessing personality disorders: implications for research and the evolution of Axis II. *American Journal of Psychiatry*, **154**, 895–903.

Zanarini, M. C., Gunderson, J. G., Frankenberg, F. R., et al (1990) Discriminating borderline personality disorder from other Axis II disorders. *American Journal of Psychiatry*, **147**, 161–167.

Zimmerman, M., Pfohl, B., Coryell, W., et al (1988) Diagnosing personality disorder in depressed patients. A comparison of patient and informant interviews. *Archives of General Psychiatry*, **45**, 733–737.