

---

# Clinical assessment of personality

Patricia R. Casey

---

The importance of personality is recognised tacitly by the development of the multiaxial classifications of ICD–10 (World Health Organization, 1992) and DSM–IV (American Psychiatric Association, 1993). The separation of axis 1 or current mental state diagnosis from axis 2, on which personality is described, is a recognition that personality is separate from other aspects of the patient's diagnostic status. Nevertheless, many clinicians are sceptical that personality can be assessed reliably and some hold that the diagnosis should be abandoned as being merely judgemental and pejorative (Lewis & Appleby, 1988). There is some basis for scepticism in respect of reliability in view of the evidence that even with operational criteria the clinical assessment of personality is not transportable, and the level of agreement between practising clinicians is much less than that achieved during field trials (Mellsop *et al*, 1982).

Before discussing the rationale for assessing personality, it is important to understand the basic properties of personality disorder:

- (a) the traits are present throughout adult life and assessment must be longitudinal;
- (b) the traits/behaviours lead to personal as well as interpersonal problems for those close to the patient, as well as for less intimate contacts (e.g. workmates);
- (c) the traits are separate and distinct from abnormalities found during episodes of illness;
- (d) the abnormal behaviours/traits are themselves consistent in a variety of situations; and
- (e) personality disorder is quantitatively different from normal variations in personality.

---

## Why assess personality?

---

Although compulsory treatment is not currently legislated for in the case of personality disorder, there

are cogent reasons for not neglecting this group of disorders either in research or in the clinical setting.

An important reason for assessing personality is the possible aetiological contribution of personality disorder to some axis 1 disorders. In particular, links with alcohol abuse, adjustment reactions and anxiety disorders have been clinically documented (Casey & Tyrer, 1990). Personality disorder also underpins some behavioural abnormalities, especially parasuicide. Other associations, between schizoid personality and schizophrenia, between borderline personality and depressive disorders and between anancastic personality and depressive disorders, are uncertain and probably reflect methodological problems in freeing personality from mental state contamination.

A second and perhaps more clinically cogent reason for making the assessment arises from the contribution personality makes to the response to treatments such as antidepressants and ECT. This finding has been described irrespective of severity of illness, with personality disorder leading to more severe symptomatology at outset and a slower rate of improvement with treatment than among patients with normal personality (Casey & Butler, 1995). Also, personality disorder is associated with greater service utilisation among those with co-occurring axis 1 disorders.

---

## Common errors

---

The poor reliability in assessing personality stems from four main sources.

---

## Confounding axes 1 and 2

---

Personality may itself change, due to environmental circumstances or psychiatric disorder. The change, however, is not permanent and personality

reverts to its former type once circumstances change. Unless there is an appreciation that the assessment is but a snapshot in time of the person's traits of personality, serious misdiagnosis may occur. For example, a person sent to prison may be compliant and well behaved for the duration of the sentence although on release he may revert to aggression, irresponsibility and volatility. The trait-state issue is a particular problem in studies which use self-rating questionnaires such as the Eysenck Personality Inventory (Eysenck & Eysenck, 1964), the Personality Disorder Questionnaire (PDQ; Hyler *et al*, 1982) and the Millon Clinical Multi-Axial Inventory (Millon, 1982); but this difficulty is not confined to psychometric measures and arises in clinical practice also.

In a clinical setting serious misrepresentation of personality may occur. The depressed person may perceive himself as being worthless, incompetent and dependent and have difficulty distinguishing his premorbid personality from his illness state (Kendell & DiScipio, 1968). Alternatively, those informants providing collateral data may have a similar problem separating trait and state (Zimmerman *et al*, 1988). Even if the clinician bases the assessment on more objective behavioural measures, a similar misunderstanding may arise since the depressed person is often slow, anxious and in need of reassurance. An even greater problem arises when serious mental illness such as schizophrenia or sometimes manic-depression supervenes in adolescence, before personality is formed. This distorts the development of adult personality and results in permanent impairment. In these circumstances, premorbid personality cannot be assessed since the observed traits, although permanent, are due to chronic illness.

It is reassuring to know, however, that the test-retest reliability of semi-structured personality interviews is satisfactory ( $\kappa=0.55$ ) when the assessment is carried out carefully and with skill, taking care to emphasise the difference between state and trait to the interviewee (Loranger *et al*, 1991). Other instruments can achieve an even higher  $\kappa$  value (Tyrer *et al*, 1983).

### *Rater bias*

The assessor's perceptions of the patient and his traits may also lead to unreliability in diagnosing personality disorder. Even when several observers assess the same trait, they may score the feature differently. One source of unreliable scoring is the scorer's view of the social desirability of the trait in question. For example, the traits 'conscientiousness' or 'introspection' may be under-scored even when they cause significant problems, if the assessor views these as highly desirable features.

Conversely, traits generally viewed as undesirable (such as pessimism, anxiousness and aggression) may be scored excessively highly. A related problem is the 'halo effect' of either liking or disliking the patient, resulting in all traits being scored in the same direction irrespective of their severity.

Stereotyping may also apply in a clinical setting; for example, women may be rated as anxious, men as aggressive etc., although the prescription of diagnostic criteria ought to eliminate this problem. Evidence for gender bias comes from the study by Mellsop *et al* (1982), who compared the level of agreement for a diagnosis of personality disorder, as well as for specific categories, using three psychiatrists who independently assessed the patients clinically. Male patients received 67% of all diagnoses of antisocial personality disorder, while females received 73 and 88%, respectively, of all diagnoses of histrionic and dependent types. Two assessors favoured the antisocial diagnosis for men, while the third preferred to make the borderline and dependent diagnoses in women.

### *Over-reliance on psychometric instruments*

A number of instruments are now available for the assessment of personality (Table 1) and there is an assumption that these will enhance the reliability of personality assessment by overcoming some of the problems outlined here. Before using such an instrument, however, it is important to be familiar with its properties as these should determine suitability. Some schedules provide dimensional measures while others adhere to the ICD-10 and/or DSM-IV categories. Some are self-report questionnaires completed by the subject or informant, whereas others are interview schedules administered by a trained interviewer. Although convenient to use, self-report questionnaires generate high rates of false positives, leading to over-diagnosis of personality disorder; some recommend that questionnaires can be used for screening purposes provided they are followed by a detailed interview, although even this is questioned by others (Zimmerman, 1994). Another distinction is between subject-only schedules and those which incorporate an informant interview.

As well as the properties described above, care must be taken in making comparisons since there is substantial evidence for poor comparability between instruments. Zimmerman & Coryell<sup>1</sup> (1990) compared the diagnoses made by the Structured Interview for DSM-III Personality Disorders (SI-D-P; Stangl *et al*, 1985) and the Personality Diagnostic Questionnaire (PDQ; Hyler *et al*, 1982) and found poor diagnostic agreement for the categories of personality disorder,

Table 1. Properties of the commonly used personality measures

Schedule	Questionnaire or interview	Dimensions or categories	Subject or informant	Duration	ICD-10 or DSM-IV
IPDE (Loranger <i>et al</i> , 1987)	Interview	Categories	Subject	150 min	Both
MCMI (Millon, 1982)	Questionnaire	Categories	Subject	30 min	DSM-IV
PDQ (Hyer <i>et al</i> , 1982)	Questionnaire	Categories	Either	30 min	DSM-IV
PAS (Tyrer & Alexander, 1979)	Interview	Both	Either	30 min	Both
SAP Mann <i>et al</i> , 1981)	Interview	Categories	Informant	10 min	Both
SCID-S (Spitzer <i>et al</i> , 1987)	Interview	Categories	Subject	90 min	DSM-IV
SID-P (Stangl <i>et al</i> , 1985)	Interview	Categories	Either	90 min	DSM-IV
EPI (Eysenck & Eysenck, 1964)	Questionnaire	Dimensions	Subject	45 min	
MMPI (Hathaway & McKinley, 1940)	Interview	Dimensions	Subject	120 min	

yielding a median  $\kappa$  value of 0.13; this increased unimpressively to 0.32 for the presence of any personality disorder. Hyler *et al* (1990) also found poor comparability between the PDQ and the Structured Clinical Interview for DSM-III (SCID) ( $\kappa=0.43$ ) and between the PDQ and the Personality Disorder Examination (IPDE) ( $\kappa=0.37$ ). However, this improved somewhat when comparisons were made using interview methods only, confirming earlier findings of the inadequacy of self-report questionnaire methods (Perry, 1992).

### *Faulty concepts*

This refers to excessive reliance on the categories derived from current classifications and also to the tendency to make a diagnosis on the basis of individual abnormal traits (i.e. using too low a threshold for making the diagnosis).

#### Over-emphasis on categories

It is often assumed that the categories of personality disorder described by current classification systems are both valid and mutually exclusive. This is not the case and categories should be viewed as hypothesised rather than empirically established syndromes. There is some evidence for the validity of the antisocial category but the validity of others has been tested with inconclusive results. For example, borderline personality disorder continues to be the subject of much investigation particularly in the USA. The continuing modifications to the categories included in ICD and DSM reflects this uncertainty: for example, schizotypal is classified in axis 2 in DSM-IV and in axis 1 in ICD-10. Passive-aggressive personality disorder was incorporated in DSM-III-R (American Psychiatric Association, 1987) but not in ICD-9 (World Health Organization, 1978), and has now been removed from DSM-IV. The anxious or avoidant personality disorder was not included in ICD-9 but has been added to ICD-10

and retained in DSM-IV. The individual categories are thus in a state of flux and decisions relating to their inclusion are derived more from potential investigative merit than from definitive scientific value. Some cynicism is therefore warranted when assigning a category of personality disorder, and there is a strong argument for merely identifying the presence or absence of personality disorder.

The co-occurrence of several personality disorders is at variance with clinical practice where there is a long-established pattern of ascribing the label which describes the most prominent category. In reality, however, the criteria for several categories can be met in most patients with an axis 2 diagnosis. A recent study found that 34 subjects received a total of 92 axis 2 diagnoses using a structured informant-based interview (Sara *et al*, 1996) and other studies (Stangl *et al*, 1985; Loranger *et al*, 1987) have found an average of between 3.75 and 4.6 personality disorder diagnoses per patient. Further studies (Widiger *et al*, 1991) have confirmed that certain types co-occur with greater than expected frequency (e.g. avoidant and dependent; borderline, histrionic and antisocial). Either the co-occurring categories have a common aetiology or the actual criteria for inclusion are over-inclusive and therefore redundant. These studies highlight the problems in using categories as if they were valid, mutually exclusive entities.

Dimensional measures have been suggested by some in preference to the categorical approach used in clinical practice (Widiger, 1987). It is argued that the lack of validity of most of the categories renders them useless and that a profile describing various aspects of the subjects' personality would be scientifically more valid and would tap into the richness of the individual personality in a way that is severely curtailed by the present categorical method. However, the use of dimensions would not confirm the presence of a personality disorder but would provide a profile on all subjects, while communication about a patient's personality using the language of dimensions would be cumbersome



compared with the categorical approach. Schedules providing dimensional measures include the Eysenck Personality Inventory and the Minnesota Multiphasic Personality Inventory.

### The threshold problem

Since each personality trait lies on a continuum from absent to severe impairment, caution must be exercised in making a diagnosis in the presence of a few traits or on the basis of traits of low severity. Thompson & Goldberg (1987) found that among 52 in-patients diagnosed clinically as having hysterical personality disorder, almost half had no recorded feature of that disorder. Among patients presenting as psychiatric emergencies, Tyrer *et al* (1994) demonstrated that the application of ICD-10 clinical criteria achieved a higher prevalence for personality disorder than did the Personality Assessment Schedule (PAS), a semi-structured diagnostic interview, thus raising the possibility that the threshold of ICD-10 was too low, or that of PAS was too high, or that the clinical diagnosis was being made without due attention to the criteria specified.

Since ICD-10 specifies social impairment as the threshold at which a diagnosis of personality disorder is made, it is likely that ICD-10 will give lower prevalence rates than DSM-IV, given the lower threshold of the latter, which requires personal distress alone. For a diagnosis such as personality disorder, carrying with it substantial stigmatising baggage, it is important to adhere to a high threshold and to err on the side of caution.

---

## The approach to assessment

---

### *Informants*

---

There are convincing data that informants report different levels of personality disorder to subjects (Tyrer & Alexander, 1979; Riso *et al*, 1994) because of lack of insight, difficulties separating traits from state, and social desirability in the response set of patients. This makes the use of informants to provide collateral data essential.

The informant must have a personal knowledge of the patient for a number of years. An employer of long standing, for example, although lacking intimate knowledge of the patient, may know much relating to the person's dealings with colleagues and those outside the immediate family. Indeed, the presence of difficulties in the workplace increases the likelihood of a diagnosis of personality disorder in view of the threshold requirement for social impairment.

There is some evidence that spouses and cohabitantes are perceived as being more reliable informants than other family members such as children, parents or siblings (Brothwell *et al*, 1992). This is somewhat surprising since it could be argued that marital loyalty or, conversely, conflict would render the information unreliable. As well as having a knowledge of the patient, the duration of acquaintance must also be considered since a short relationship will be associated with exposure to a restricted number of traits. However, no particular time cut-off has been established and duration is mentioned here for consideration rather than as a definitive requirement. Among informants there is some evidence that females are perceived as being more reliable than males (Brothwell *et al*, 1992).

These findings would suggest that when possible a female should be chosen as the informant, a spouse or cohabitee in preference to other first-degree relatives, and a person who has long-standing knowledge of the patient. Even when these requirements are met, the insight and psychological awareness of the individual informant is likely to be the most important determinant of reliability. If disagreement arises between subject and informant data, Tyrer & Alexander (1979) suggest opting for the latter on the basis of slightly higher test-retest reliability.

### *Which traits or behaviours?*

---

This question focuses on two distinct approaches to personality assessment. Should the emphasis be on behaviours only or exclusively on traits? The current assessment of personality, as well as the ICD-10 and DSM-IV classifications, are based on the measurement of traits although this method is of questionable reliability. By using measures of behaviour ('behavioural exemplars') to represent traits the problem of lack of insight is overcome. Moreover, behaviours are more easily observable and quantifiable than the more nebulous traits commonly used in practice. The greater reliability of objective measures such as behavioural exemplars has been confirmed in a study by Mellsop *et al* (1982) in which reliability between three clinical psychiatrists was highest for anti-social personality disorder, with its reliance on objective measures to make the diagnosis. In spite of having advocates (Livesley, 1986) the use of behavioural measures is not well developed and would require the development of multiple behavioural exemplars since single measures could arise from other traits as well as from other sources and would provide only a limited sample of the construct. This limitation was demonstrated by

Thompson & Goldberg (1987) who found that the specific behaviours were readily attributed to personality disorder rather than to other possible causes such as depressive illness.

The trait approach is therefore currently used in assessing personality and begs the question of which traits should be assessed. The possible list is endless but guidance comes from the studies involved in developing the PAS (Tyrer & Alexander, 1979). The first task was to identify traits that have been described in a number of definitions of personality disorder (including ICD and DSM). During field studies those that were infrequent or of insufficient severity were excluded. Subsequently, factor analysis was used to shorten the list further by examining the degree of overlap. Finally, 24 traits were included (Table 2), since they were found to be sufficiently independent and quantifiable. By focusing on these traits during clinical assessment, the doctor will avoid the danger of making diagnoses on the basis of overlapping traits, some of which may not be relevant to making the diagnosis.

The dependence on traits exclusively may be viewed as problematic but by eliciting examples of behavioural abnormalities and encouraging the use of anecdotes to illustrate these abnormal traits a consistent pattern can be identified (Luborsky, 1990).

Table 2. Traits investigated in the development of the PAS (Tyrer & Alexander, 1979)

Traits excluded	Traits included
Possessiveness	Pessimism
Dominance	Worthlessness
Fickleness	Optimism
Ambivalence	Lability
Disorderliness	Anxiousness
Passiveness	Suspiciousness
Stubbornness	Introspection
Humility	Shyness
Anger	Aloofness
Insensitivity	Sensitivity
Weakness	Vulnerability
Fanaticism	Irritability
Gullibility	Impulsiveness
Foolhardiness	Aggression
Conformity	Callousness
Timidity	Irresponsibility
Carelessness	Childishness
Vanity	Resourcelessness
Grandiosity	Dependence
Touchiness	Submissiveness
Shallowness	Conscientiousness
	Rigidity
	Eccentricity
	Hypochondriasis

## The LEAD standard

Longitudinal Expert Evaluation using All Data (LEAD) was first described by Spitzer (1983) in the context of a controversy about whether personality should standardly be assessed by interviewer-administered schedule or by a clinician's interview. Spitzer argued that in the absence of an infallible diagnosis, the best diagnosis was arrived at by an expert on the disorder observing the patient over a period of time and using information from many sources including other hospital records, collateral from family and friends and data from all treating staff who would provide information on maladaptive behaviours and traits in various contexts and independently of episodes of illness. This 'state of the art' diagnosis could be a base from which other techniques, including interview schedules, could be compared and validated.

Although experimental, this approach has won some accolades (Skodol *et al*, 1991). Difficulties associated with LEAD include the time involved in making a diagnosis, its adherence to diagnostic criteria derived from the (imperfect) current classifications, and the untested reliability of the LEAD procedure. However, the rigour of the method is such that many of the common errors in assessment described above are eliminated or minimised.

## Conclusions

Numerous schedules are used in making personality assessments. The approach in clinical practice and the commonly used interview schedules are often inadequate and unreliable. Undue reliance on the categories listed in the current classifications is also problematic. Distinction between state and traits, use of informants, and a high threshold for diagnosis improve reliability. There is no 'gold standard' for the diagnosis of personality disorder although the LEAD approach is currently the most promising.

## References

- American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn, revised) (DSM-III-R). Washington, DC: APA.
- (1993) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). Washington, DC: APA.
- Brothwell, J., Casey, P. & Tyrer, P. (1992) Who gives the most reliable account of a psychiatric patient's personality? *Irish Journal of Psychological Medicine*, 9, 90–93.

- Casey, P. & Tyrer, P. (1990) Personality disorder and psychiatric illness in general practice. *British Journal of Psychiatry*, **156**, 261–265.
- & Butler, E. (1995) The effects of personality on response to ECT in major depression. *Journal of Personality Disorders*, **9**, 134–142.
- Eysenck, H. J. & Eysenck, S. B. G. (1964) *Manual of the Eysenck Personality Inventory (EPQ)*. London: University of London Press.
- Hathaway, S. R. & McKinley, J. C. (1940) A multiphasic personality schedule (Minnesota): Construction of the schedule. *Journal of Psychology*, **10**, 249–254.
- Hyler, S., Reider, R., Spitzer, R., et al (1982) *Personality Diagnostic Questionnaire*. New York: New York State Psychiatric Institute.
- , Skodol, A. E., Kellman, H. D., et al (1990) Validity of the Personality Diagnostic Questionnaire – Revised: comparison with two structured interviews. *American Journal of Psychiatry*, **147**, 1043–1048.
- Kendell, R. E. & DiScipio, W. J. (1968) Eysenck Personality Inventory scores of patients with depressive illness. *British Journal of Psychiatry*, **114**, 767–770.
- Lewis, G. & Appleby, L. (1988) Personality disorder: the patients psychiatrists dislike. *British Journal of Psychiatry*, **153**, 44–49.
- Livesley, W. J. (1986) Trait and behavioural prototypes of personality disorders. *American Journal of Psychiatry*, **143**, 728–732.
- Loranger, A. W., Susman, V. L., Oldham, M. M., et al (1987) The Personality Disorder Examination: a preliminary report. *Journal of Personality Disorders*, **1**, 1–13.
- , Lenzenweger, M. F., Gartner, A. F., et al (1991) Trait–state artifacts and the diagnosis of personality disorders. *Archives of General Psychiatry*, **48**, 720–728.
- Luborsky, L. & Crits-Christoph, P. (1990) *Understanding Transference: The CCRT Method*. New York, Basic Books.
- Mann, A. H., Jenkins, R., Cutting, J. C., et al (1981) The development and use of a standardized assessment of abnormal personality. *Psychological Medicine*, **11**, 839–847.
- Mellsop, G., Varghese, F., Joshua, S., et al (1982) The reliability of axis II of DSM–III. *American Journal of Psychiatry*, **139**, 1360–1361.
- Millon, T. (1982) *Millon Clinical Multi-Axial Inventory* (2nd edn). Minneapolis, MN: Interpretive Scoring Systems.
- Perry, J. C. (1992) Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry*, **149**, 1645–1653.
- 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.
- Sara, G., Raven, P. & Mann, A. (1996) A comparison of DSM–III–R and ICD–10 personality disorder criteria in an outpatient population. *Psychological Medicine*, **26**, 151–160.
- Skodol, A. E., Rosnick, L., Kellman, D., et al (1991) Development of a procedure for validating structured assessments of axis II. In *Personality Disorders: New Perspectives on Diagnostic Validity* (ed. J. M. Oldham). Washington, DC: American Psychiatric Press.
- Spitzer, R. L. (1983) Psychiatric diagnosis: are clinicians still necessary? *Comprehensive Psychiatry*, **24**, 399–411.
- , Williams, J. B. W. & Gibbons, M. (1987) *Structured Clinical Interview for DSM–III–R, Patient Version*. New York: New York State Psychiatric Institute.
- Stangl, D., Pfohl, B., Zimmerman, M., et al (1985) Structured interview for DSM–III personality disorders. *Archives of General Psychiatry*, **42**, 591–596.
- Thompson, D. J. & Goldberg, D. (1987) Hysterical personality disorder. The process of diagnosis in clinical and experimental settings. *British Journal of Psychiatry*, **150**, 241–245.
- Tyrer, P. & 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.
- , Merson, S., Onyett, S., et al (1994) The effects of personality disorder on clinical outcome, social networks and adjustment: a controlled clinical trial of psychiatric emergencies. *Psychological Medicine*, **24**, 731–740.
- Widiger, T. A. (1987) The multidimensional scaling of the DSM–III personality disorders. *Acta Psychiatrica Scandinavica*, **44**, 557–563.
- , Frances, A. J., Harris, M., et al (1991) Co-morbidity among axis II disorders. In *Personality Disorder: New Perspectives on*

- Diagnostic Validity* (ed. J. M. Oldham). Washington, DC: American Psychiatric Press.
- World Health Organization (1978) *Mental Disorders. Glossary and Guide to their Classification in Accordance with the 9th Revision of the International Classification of Diseases (ICD–9)*. Geneva: WHO.
- (1992) *The ICD–10 Classification of Mental and Behavioural Disorders*. Geneva: WHO.
- Zimmerman, M. (1994) Diagnosing personality disorders: a review of issues and research methods. *Archives of General Psychiatry*, **51**, 225–245.
- & Coryell, W. H. (1990) Diagnosing personality disorders in the community: a comparison of self-report and interview measures. *Archives of General Psychiatry*, **47**, 527–531.
- , 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.

## Multiple choice questions

- Personality can be reliably assessed by using:
  - assessment during an episode of illness
  - informants who are close to the patient
  - information from short-term acquaintances
  - a questionnaire.
- Categories of personality disorder:
  - are well validated
  - have high test–retest reliability
  - are best described dimensionally
  - rarely co-occur with other categories.
- Psychometric assessment of personality:
  - is rapidly carried out
  - is more reliable than clinical assessment
  - depends for reliability on the properties of the schedule
  - ideally utilises information from informants.
- Ideally, personality disorder is assessed using:
  - traits
  - categories
  - behavioural exemplars
  - multiple sources of information.
- The trait approach to personality assessment:
  - is the most commonly used method
  - suffers from the problem of trait overlap
  - requires a degree of insight in the interviewee
  - is the basis for the ICD–10 classification.

### MCQ answers

1	2	3	4	5
a F	a F	a F	a F	a T
b T	b F	b F	b F	b T
c F	c F	c T	c T	c T
d F	d F	d T	d T	d T