

Reporting of personality disorder symptoms in a forensic inpatient sample: Effects of mode of assessment and response style

Published in *The Journal of Forensic Psychiatry & Psychology*, September 2006; 17(3): 431 – 441

DOI: 10.1080/14789940600775436

Authors: SEAN WHYTE¹, SIMONE FOX², & ADRIAN COXELL³

¹Three Bridges Regional Secure Unit, West London Mental Health NHS Trust, Uxbridge Road, Southall, Middlesex

²Wells Unit, Adolescent Forensic Service, West London Mental Health NHS Trust

³Chartered Clinical Psychologist, Broadmoor Hospital, Berkshire

Abstract

Self-report measures of personality disorder are known to over-report personality disorder traits consistently relative to measures based on a clinical interview. This study tests the hypothesis that there is a relationship between a participant's tendency to give socially desirable responses and the discrepancy between their results on self-report and interview-based measures. The relative over-reporting of the self-report measure, the PDQ4, was confirmed, but no association was found between the magnitude of the over-reporting and the tendency to give socially desirable responses.

Keywords: Personality disorder, assessment, social desirability bias

INTRODUCTION

Personality disorders are psychiatric disorders characterized by chronic patterns of behaviour that are inflexible and present across a broad range of settings and situations (Ward, 2004). A diagnosis of personality disorder can be associated with a number of negative outcomes including an increased risk of suicide (Moran et al., 2003a) and a decreased likelihood of benefiting from psychological therapy for Axis I disorder (Tyrer, Gunderson, Lyons, & Tohen, 1997). Personality disorders are also associated with an increased risk of criminal behaviours (Moran et al., 2003b) including homicide (Joyal, Putkonen, Paavola, & Tiihonen, 2004). The assessment and diagnosis of personality disorder is becoming an increasingly important issue due to proposed changes in mental health legislation. For instance, in the United Kingdom, the government's draft Mental Health Bill (Department of Health, 2005) regards personality disorder as on a par with other mental disorders, rather than in the separate category of "psychopathic disorder", with its own rules for diagnosis and treatment. It allows for persons with mental disorder, which is very broadly defined, to be detained if they present a "substantial risk of serious harm to other persons". Because of their previous exclusion from many mental health services, there is currently limited expertise or

experience in the assessment and treatment of dangerous persons with personality disorder (Maden & Tyrer, 2003).

Personality disorder may be assessed in a variety of different ways, including clinical interview, interview and review of the participant's notes, interview of both participant and an informant, or using self-report measures. There is now a wide variety of instruments measuring the main categories of personality disorder, particularly using the DSM classification system: examples include the Millon Clinical Multiaxial Inventory (Millon, 1984), the Minnesota Multiphasic Personality Inventory (Butcher, Dahlstrom, Graham, Tellegen, & Kraemmer, 1989), the International Personality Disorder Examination (Loranger et al., 1994), the Personality Assessment Schedule (Tyrer, Alexander, & Ferguson, 2000), and the Personality Diagnostic Questionnaire (Hyler, 1994). As yet, no 'gold standard' exists for the diagnosis of personality disorder, and concordance between different structured interviews (Oldham, Skodol, Kellman, & Hyler, 1992) and between structured interviews and self-report measures of personality disorder (Blackburn, Donnelly, Logan, & Renwick, 2004; De Ruiter & Greeven, 2000; Schotte et al., 2004) is often low. Typically, self-report measures have also been shown to over-diagnose personality disorders when compared to structured interviews (Bodlund, Grann, Ottoman, & Svanborg, 1998; Hyler, Skodol, Kellman, Oldham, & Rosnick, 1990).

One reason for this low concordance could be response bias. Response biases include the under-reporting by participants of negatively-perceived symptoms and the over-reporting of positive traits and behaviours (i.e., socially desirable responding) and, conversely, the over-reporting of symptoms (i.e., malingering). Measures of socially desirable responding have been found to correlate negatively with self-reports of symptoms of mental disorder: for example, McNiel, Eisner, and Binder (2000) found this with command auditory hallucinations in psychiatric inpatients, and Tanaka and Kameoka (1986) found this with anxiety and depression in university students. This is an important finding since it suggests that reporting of symptoms may be affected by social desirability even when participation is anonymous – that is, even when participants' responses are unknown and cannot lead to negative consequences.

Social desirability can also exert an effect on measures relating to personality. For example, Tourangeau, Rips, and Rasinski (2000) found a higher rate of reporting of 'sensitive behaviours' when using a questionnaire as compared to an interview. Krahe (1989) found that research participants are able to manipulate their scores convincingly on scales from a personality test consistent with pre-testing instructions to respond in a given way.

It has been argued that socially desirable responding may be more evident in forensic contexts where reporting of various traits or behaviours may result in an increased likelihood of detention in hospital or prison, and authors have therefore advised against the use of self-report inventories without robust measures of distortion when assessing forensic populations (e.g., Hare, Hart, & Harpur, 1991). What is not known at this time, however, is whether male forensic inpatients' responses to questions about personality disorder symptoms differ between self-report and interview-based measures on the basis of such socially desirable responding. This study was therefore designed to investigate whether the discrepancy found in other studies between self-report and interview-based measures of

personality disorder also occurs in our forensic male inpatient sample; and to test the hypothesis that this discrepancy is associated with each participant's tendency to give socially desirable responses.

METHOD

Participants

Participants were recruited from across all medium- and low-secure wards of the inpatient forensic service of the West London Mental Health NHS Trust. This study population was chosen for the anticipated high prevalence of abnormal personality traits. Permission was sought from each Responsible Medical Officer (RMO) to approach their patients; no RMO refused permission for their patients to be considered as study participants. Male patients, irrespective of diagnosis, who were deemed capable of giving informed consent and who had at least a basic command of English were invited to take part. A set of power calculations indicated that a sample of 80 participants would have at least 95% power to detect the anticipated discrepancy between interview- and questionnaire-based assessments at $p \leq .05$. A total of 146 participants were invited to participate: 55 refused consent, and 91 gave valid informed consent. Of these 91, 11 were excluded before completing the study (because they withdrew consent or, for example, because their English turned out to be too poor for the tasks), leaving 80 participants. The mean age of participants was 38 (range 21 – 60).

Instruments

The Personality Assessment Schedule (PAS; Tyrer et al., 2000) was used as the interview and informant-based measure of personality disorder. The PAS is a semi-structured interview schedule which has been shown to be reliable and valid (Tyrer & Alexander, 1979; Tyrer, Alexander, & Cicchetti, 1979). A version that uses the diagnostic categories of the DSM-IV was employed.

The Personality Diagnostic Questionnaire, DSM-IV version (PDQ4) was used as the self-report measure of personality disorder. The PDQ4 is a 99-item questionnaire which screens for the presence of DSM Axis II personality disorders and has been shown to be both reliable and valid (Dubro, Wetzler, & Kahn, 1988; Hyler et al., 1988, 1990; Pfohl, Coryell, Zimmerman, & Stangl, 1987).

The Balanced Inventory of Desirable Reporting, seventh edition (BIDR7; subsequently renamed the Paulhus Deception Scales; Paulhus, 1998), was used to measure participants' tendency to give socially desirable responses. This is a 40-item self-report inventory that measures an individual's tendency to give socially desirable responses. Two principal and relatively independent subscales are reported: self-deceptive enhancement (SDE), which represents 'an unconscious favourability bias closely related to narcissism and which taps self-deception in the sense of a pervasive lack of insight'; and impression management (IM), which represents 'a well-known category of social desirability measures aimed at the crude form of dissimulation known as faking or lying' (pp. 144).

Procedure

All consenting participants, and their informants (usually their primary nurse), were interviewed in person by the psychiatrist SW, using the PAS. Participants were also seen by SF, AC (both clinical psychologists), or a research assistant, and given the PDQ4 and BIDR7 to complete. The two assessments were conducted at least seven days apart, to minimize bias caused by memory of earlier questions. The order of questionnaire administration and interview was counterbalanced.

Data analysis

Data were entered into Microsoft Excel 2002 and analysed using Microsoft Excel Analysis ToolPak. Scores on both the PAS and the PDQ4 for each personality disorder category were summed, so as to give aggregate scores for the three DSM-IV clusters and a total score. The categories used were paranoid, schizoid, and schizotypal (Cluster A), histrionic, antisocial, borderline, and narcissistic (Cluster B), and avoidant, dependent, and obsessive-compulsive (Cluster C). These aggregate scores were then scaled to give scores from 0 to 10 for each of the clusters and from 0 to 30 for the totals. References to scores in the Results section are to these scaled scores.

Initial analysis demonstrated that there were no significant outliers, and that the PAS and PDQ4 scores were both normally distributed. The student's t test was therefore used for analysing the PAS and PDQ4 scaled scores. The Wilcoxon signed ranks test was used when comparing numbers of diagnoses made by an instrument (i.e., the number of categories in which the raw score was above the cut-off score for a diagnosis of personality disorder in that category).

The kappa (k) test was used for testing agreement on diagnosis between the two measures, as percentage concordances will tend to overestimate agreement because of the high probability of agreement by chance. Conventionally, k values of less than .2 are regarded as demonstrating poor agreement; .2 – .4 is 'fair'; .4 – .6 is 'moderate'; .6 – .8 is 'good'; and values over 0.8 indicate very good agreement.

Further inspection demonstrated an approximate straight-line relationship between the major variables, and the Pearson product-moment correlation coefficient r was therefore used to analyse correlations between variables. Conventionally (see, e.g., Aron & Aron, 1994; Kenney & Keeping, 1962), r values of less than +.3 indicate little or no association, values from +.3 to +.7 indicate weak correlations, and values above +.7 indicate a strong correlation.

Results

Examination of the individual pairs of results showed that the PDQ4 total scores were, on average, 15.6% greater than the corresponding PAS total scores; this average difference was consistent across the three clusters (Cluster A mean difference 16.6%, Cluster B 13.6%, Cluster C 16.7%). The mean total score on the PAS was 6.67 out of 30 whereas the mean total score on the PDQ4 was 11.37 out of 30; this difference was statistically significant

($t_{1710.29}$, df_{178} , $p_{5.0001}$). Significant differences also existed between the PAS and PDQ4 scores for each of the clusters (data not shown).

An analysis of the number of diagnoses of personality disorder made for each participant by each instrument was also performed, in order to take into account the different cut-off scores used by the PAS and PDQ4. The mean number of diagnoses of personality disorder per participant was 2.3 for the PAS and 3.2 for the PDQ4; this difference was also statistically significant (W_{1773} , $p_{5.01}$), as were those for Clusters A and C (see Table I).

Agreement between the two instruments on whether individual personality disorders were present was mostly poor or fair, as shown in Table II. Agreement varied from 50% ($k_{.05}$) for obsessive-compulsive personality disorder to 84% ($k_{.48}$) for dependent personality disorder. A weak but significant correlation ($r_{.60}$, $p_{5.00001}$) was found between the total PAS scores and the total PDQ4 scores; the scatter plot in Figure 1 demonstrates the same correlation. This correlation also held for each of the three clusters (Cluster A $r_{.51}$, $p_{5.00001}$; Cluster B $r_{.48}$, $p_{5.00001}$; Cluster C $r_{.48}$, $p_{5.0001}$). Correlations at the level of individual diagnoses ranged from $r_{.25}$ to $r_{.49}$, all of which were statistically significant.

Table I. Mean numbers of personality disorder diagnoses.

	Self-report (PDQ)	Interview (PAS)	Difference
Cluster A	1.1	0.7	$p < .01$
Cluster B	1.1	0.9	$p = NS$
Cluster C	1.1	0.7	$p < .01$
Total	3.2	2.3	$p < .01$

Table II. Agreement between PAS and PDQ4 for each personality disorder category and cluster.

Category or cluster	Percentage concordance	Kappa
Paranoid	61.3%	.24
Schizotypal	68.8%	.21
Schizoid	67.5%	.03
Histrionic	70.0%	.12
Narcissistic	68.8%	.18
Borderline	72.5%	.15
Antisocial	71.3%	.30
Avoidant	62.5%	.14
Dependent	83.8%	.48
Obsessive-compulsive	50.0%	.05
Cluster A	55.0%	.09
Cluster B	62.5%	.24
Cluster C	58.8%	.22
Any personality disorder	68.8%	.19

The mean scores (and standard deviations) for impression management, self-deceptive enhancement, and BIDR total were 8.1 (3.2), 5.6 (4.2), and 13.7 (6.0). All were considerably greater than published figures for the general population, 6.7 (4.0), 2.2 (2.3),

and 8.9 (3.7), and for prison entrants, 5.3 (3.6), 2.2 (2.7), and 7.5 (3.5) (Endler & Parker, 2005).

There were consistent negative correlations between the PAS scores and the total BIDR score, and also between the PDQ scores and the total BIDR score, as shown in Table III. Approximately half of the correlations were statistically significant ($p < .05$). There was a trend for the correlations between PAS and BIDR scores to be more negative than those between PDQ and BIDR scores. However, the magnitude of the correlations was very small: the largest was that between the score for antisocial personality disorder on the PDQ and the BIDR7 score (70.304, $p < .01$), and the rest were all in the 'little or no association' range $0.0 - +0.3$. This was also the case when the cut-off score was used with the BIDR7 data and the correlations were re-examined using the phi coefficient for binary data (data not shown).

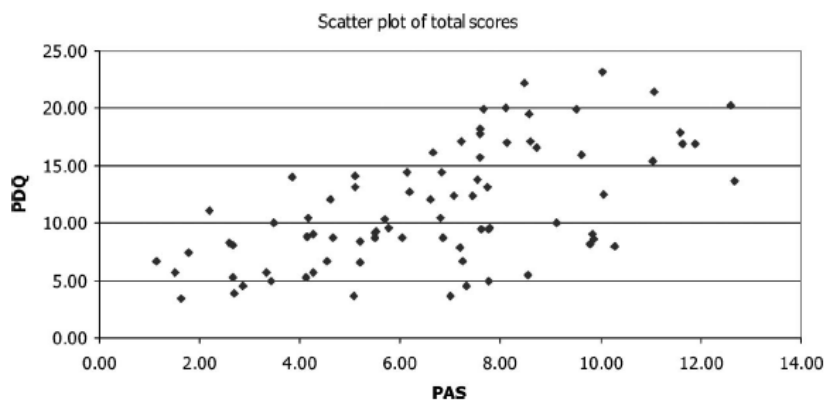


Figure 1. Scatter plot of total PAS and PDQ scores.

Table III. Correlations between desirable reporting and PAS or PDQ4 core.

Pearson's r (p)	Self-report (PDQ)	Interview (PAS)
Paranoid	-.099	-.220*
Schizotypal	-.110	-.131
Schizoid	+.034	-.193*
Histrionic	-.035	-.229*
Narcissistic	-.062	-.164
Borderline	-.209*	-.255*
Antisocial	-.304**	-.260**
Avoidant	-.233*	-.177
Dependent	-.115	-.156
Obsessive-compulsive	-.161	-.157
Cluster A	-.125	-.168
Cluster B	-.239*	-.194
Cluster C	-.269*	-.156
Total PD symptom score	-.218*	-.199

*Significant at $p < .05$.

**Significant at $p < .01$.

The differences between the PAS and PDQ4 scores for each participant were then correlated with their impression management (IM), self-deceptive enhancement (SDE), and total BIDR scores. These figures for signed differences are displayed in Table IV. No statistically significant correlation was found between any of the PAS –PDQ4 differences and IM, SDE, or total BIDR scores. The same lack of any correlation was also found when the absolute

differences – that is, the magnitude of the difference in each case, discarding the sign of the difference – in PAS and PDQ4 scores were correlated with IM, SDE, and total BIDR scores

Table IV. Correlations between PAS–PDQ4 signed difference and desirable reporting.

Pearson's <i>r</i> (<i>p</i>)	Impression management	Self-deceptive enhancement	Total BIDR score
Cluster A	–.16 (NS)	.06 (NS)	–.04 (NS)
Cluster B	–.19 (NS)	.07 (NS)	–.05 (NS)
Cluster C	–.18 (NS)	–.16 (NS)	–.21 (NS)
Total	–.22 (<i>p</i> < .05)	–.01 (NS)	–.13 (NS)

(data not shown)

DISCUSSION

This study has confirmed the earlier finding (e.g., Zimmerman & Coryell, 1990) that self-report measures such as the PDQ4 yield higher rates of personality disorder symptomatology than structured interviews. This is the case both in terms of categorical diagnoses (the PDQ4 diagnosed, on average, 0.9 times more disorders per patient than the PAS) and in terms of dimensional score, which when scaled for comparison was on average 16% higher with the PDQ4. The short-term temporal instability of personality (given that the assessments were conducted a week apart) does not appear to have affected this finding.

A weakness of this study, to take into account when assessing correlations with impression management, is that the participants were already in hospital: that is, they might no longer have had the strong incentive to manage others' impressions of them that they would have had before trial and admission. For example, Ahlmeyer, Heil, McKee, and English (2000) found that sexual offenders admitted to dramatically greater numbers of offences after trial than before trial. Against this, however, other authors have found that 'other-deception', presumed to be related to social desirability, persists in patients in secure psychiatric hospitals (e.g., Gudjonsson & Moore, 2001).

The correlation between the PDQ4 total score and the PAS total score was significant but weak, as were the correlations for each cluster ($r = .48 - .60$, $p = .0001 - .0001$). When the two instruments were compared using diagnostic cut-offs, agreement was 'fair' at best ($k = .48$) and 'poor' at worst ($k = .05$) for individual diagnoses; the kappa scores for comparisons at the level of the cluster and at the level of 'any disorder' were in the lower half of the same range. One interpretation of this is that while both instruments are measuring the same underlying personality disorder traits in the individual patient (and are therefore correlated with each other), they are both imperfect measures, and imperfect in different ways, so that their agreement is poor. It may also be the case that at least some of the individual personality disorder diagnoses which these instruments seek to assess are not, as currently defined, valid clinical entities (for a further discussion of this point see, e.g., Tyrer, 2001).

This study set out to test the hypothesis that the difference between individual patients' self-report and interview-based scores would be explained by each patient's tendency to give socially desirable responses. No clinically significant association was found between the difference in PAS and PDO4 scores and socially desirable reporting (IM and SDE) as measured by the BIDR7. Given that a power calculation had indicated that the number of participants was sufficient to demonstrate any clinically significant association which might exist, it was concluded that socially desirable reporting as measured by the BIDR7 does not, at least in this study, explain the discrepancy between self-report and interview-based measures of personality disorder.

ACKNOWLEDGEMENTS

We are grateful to our research assistants Ligia Costa and Elizabeth Nyakapanka, who administered a number of questionnaires to participants, and to Paul Bassett of the Statistics Consultancy, who advised us on the statistical analysis.

REFERENCES

- Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraphy on admissions of victims and offences in adult sexual offenders. *Sexual Abuse: Journal of Research and Treatment*, 12, 123–138.
- Aron, A., & Aron, E. (1994). *Statistics for psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Blackburn, R., Donnelly, J. P., Logan, C., & Renwick, S. J. D. (2004). Convergent and discriminative validity of interview and questionnaire measures of personality disorder in mentally disordered offenders: A multitrait-multimethod analysis using confirmatory factor analysis. *Journal of Personality Disorders*, 18, 129–150.
- Bodlund, O., Grann, M., Ottoman, H., & Svanborg, C. (1998). Validation of the self-report questionnaire DIP-Q in diagnosing DSM-IV personality disorders: A comparison of three psychiatric samples. *Acta Psychiatrica Scandinavica*, 97, 433–439.
- Butcher, J., Dahlstrom, W., Graham, J., Tellegen, A., & Kraemmer, B. (1989). *Minnesota Multiphasic Personality Inventory (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Department of Health. (2005). *Draft mental health bill*. Retrieved January 5, 2005, from http://www.dh.gov.uk/policyandguidance/healthandsocialcaretopics/mentalhealth/mentalhealtharticle/fs/en?content_id¼4089593&chk¼t9p0rz
- De Ruiter, C., & Greeven, P. G. J. (2000). Personality disorders in a Dutch forensic psychiatric sample: Convergence of interview and self-report measures. *Journal of Personality Disorders*, 14, 162–170.
- Dubro, A. F., Wetzler, S., & Kahn, M. W. (1988). A comparison of three self-report questionnaires for the diagnosis of DSM-III personality disorders. *Journal of Personality Disorders*, 2, 256–266.

- Endler, N. S., & Parker, J. D. A. (2005). *PDS Paulhus Deception Scales: Technical brochure*. Toronto/Buffalo: Multi-Health Systems Inc.
- Gudjonsson, G. H., & Moore, E. (2001). Self-deception and other-deception among admissions to a maximum security hospital and a medium secure unit. *Psychology, Crime and Law*, 7, 25–31.
- Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, 100, 391–398.
- Hyer, S. E. (1994). *Personality Diagnostic Questionnaire – 4*. New York: New York State Psychiatric Institute.
- Hyer, S. E., Rieder, R. O., Williams, J. B.W., Spitzer, R. L., Hendler, J., & Lyons, M. (1988). The Personality Diagnostic Questionnaire: Development and preliminary results. *Journal of Personality Disorders*, 2, 229–237.
- Hyer, S. E., Skodol, A. E., Kellman, H. D., Oldham, J., & Rosnick, L. (1990). The validity of the Personality Diagnostic Questionnaire: A comparison with two structured interviews. *American Journal of Psychiatry*, 147, 1043–1048.
- Joyal, C. C., Putkonen, A., Paavola, P., & Tiihonen, J. (2004). Characteristics and circumstances of homicidal acts committed by offenders with schizophrenia. *Psychological Medicine*, 34, 433–442.
- Kenney, J. F., & Keeping, E. S. (1962). Linear regression and correlation. In *Mathematics of statistics* (3rd ed., pp. 252–285). Princeton, NJ: Van Nostrand.
- Krahe´, B. (1989). Faking personality profiles on a standard personality inventory. *Personality and Individual Differences*, 10, 437–443.
- Loranger, A. W., Sartorius, N., Andreoli, A., Berger, P., Bruckheim, P., & Charnabasavanna, S. M., et al. (1994). The International Personality Disorder Examination: The WHO/ ADAMHA international pilot study of personality disorders. *Archives of General Psychiatry*, 51, 212–224.
- Maden, T., & Tyrer, P. (2003). Dangerous and severe personality disorders: A new personality concept from the United Kingdom. *Journal of Personality Disorders*, 17, 489–496.
- McNiel, D. E., Eisner, J. P., & Binder, R. L. (2000). The relationship between command hallucinations and violence. *Psychiatric Services*, 51, 1288–1292.
- Millon, T. (1984). *The Millon Clinical Multiaxial Inventory*. Minneapolis: National Computer Systems.
- Moran, P., Walsh, E., Tyrer, P., Burns, T., Creed, F., & Fahy, T. (2003a). Does comorbid personality disorder increase the risk of suicidal behaviour in psychosis? *Acta Psychiatrica Scandinavica*, 107, 441–448.
- Moran, P., Walsh, E., Tyrer, P., Burns, T., Creed, F., & Fahy, T. (2003b). Impact of comorbid personality disorder on violence in psychosis: Report from the UK700 trial. *British Journal of Psychiatry*, 182, 129–134.

- Oldham, J.M., Skodol, A. E., Kellman, H. D., & Hyler, S. E. (1992). Diagnosis of DSM-III-R personality disorders by two structured interviews: Patterns of comorbidity. *American Journal of Psychiatry*, 149, 213–220.
- Paulhus, D. L. (1998). *Manual for the Balanced Inventory of Desirable Responding: Version 7*. Toronto/Buffalo: Multi-Health Systems.
- Pfohl, B., Coryell, W., Zimmerman, M., & Stangl, D. (1987). Prognostic validity of self-report and interview measures of personality disorder in depressed in-patients. *Journal of Clinical Psychiatry*, 48, 469–472.
- Schotte, C. K. W., De Doncker, D. A. M., Dmitruk, D., Van Mulders, I., De Haenen, H., & Cosyns P. (2004). The ADP-IV questionnaire: Differential validity and concordance with the semi-structured interview. *Journal of Personality Disorders*, 18, 405–419.
- Tanaka, M. J., & Kameoka, V. A. (1986). Reliabilities and concurrent validities of popular selfreport measures of depression, anxiety, and social desirability. *Journal of Consulting and Clinical Psychology*, 54, 328–333.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.
- Tyrer, P. (2001). Personality disorder. *British Journal of Psychiatry*, 179, 81–84.
- Tyrer, P., & Alexander, J. (1979). Classification of personality disorder. *British Journal of Psychiatry*, 135, 163–167.
- Tyrer, P., Alexander, M. S., & Cicchetti, D. (1979). Reliability of a schedule for rating personality disorders. *British Journal of Psychiatry*, 135, 168–174.
- Tyrer, P., Alexander, J., & Ferguson, B. (2000). Personality Assessment Schedule. In P. Tyrer (Ed.), *Personality disorders: Diagnosis, management and course* (2nd ed., pp. 133–159). Oxford: Butterworth-Heinemann.
- Tyrer, P., Gunderson, J., Lyons, M., & Tohen, M. (1997). Extent of comorbidity between mental state and personality disorders. *Journal of Personality Disorders*, 11, 242–259.
- Ward, R. K. (2004). Assessment and management of personality disorders. *American Family Physician*, 70, 1505–1512.
- Zimmerman, M., & 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.