July 2007

Comparison of alternative models for personality disorders

Leslie C. Morey
Texas A & M University - College Station

Christopher J. Hopwood
Texas A & M University - College Station

John G. Gunderson
McLean Hospital, Harvard Medical School

Andrew E. Skodol
New York State Psychiatric Institute, Columbia University

M. Tracie Shea
Brown University

See next page for additional authors

Recommended Citation

Morey, L. C., Hopwood, C. J., Gunderson, J. G., Skodol, A. E., Shea, M. T., Yen, S., Stout, R. L., Zanarini, M. C., Grilo, C. M., Sanislow, C. A., & McGlashan, T. H. (2007). Comparison of alternate models for personality disorders. *Psychological Medicine, 37(7)*, 983-994.

Psychological Medicine, 2007, **37**, 983–994. © 2006 Cambridge University Press
doi:10.1017/S0033291706009482 First published online 23 November 2006 Printed in the United Kingdom

Comparison of alternative models for personality disorders

LESLIE C. MOREY[1]*, CHRISTOPHER J. HOPWOOD[1], JOHN G. GUNDERSON[2], ANDREW E. SKODOL[3], M. TRACIE SHEA[4], SHIRLEY YEN[4], ROBERT L. STOUT[5], MARY C. ZANARINI[2], CARLOS M. GRILO[6], CHARLES A. SANISLOW[6] AND THOMAS H. McGLASHAN[6]

[1] *Texas A&M University, College Station, TX, USA*; [2] *McLean Hospital, Harvard Medical School, MA, USA*; [3] *New York State Psychiatric Institute, Columbia University, NY, USA*; [4] *Brown University, RI, USA*; [5] *Decision Sciences Institute, RI, USA*; [6] *Yale University School of Medicine, CT, USA*

ABSTRACT

Background. The categorical classification system for personality disorder (PD) has been frequently criticized and several alternative dimensional models have been proposed.

Method. Antecedent, concurrent and predictive markers of construct validity were examined for three models of PDs: the Five-Factor Model (FFM), the Schedule for Nonadaptive and Adaptive Personality (SNAP) model and the DSM-IV in the Collaborative Study of Personality Disorders (CLPS) sample.

Results. All models showed substantial validity across a variety of marker variables over time. Dimensional models (including dimensionalized DSM-IV) consistently outperformed the conventional categorical diagnosis in predicting external variables, such as subsequent suicidal gestures and hospitalizations. FFM facets failed to improve upon the validity of higher-order factors upon cross-validation. Data demonstrated the importance of both stable trait and dynamic psychopathological influences in predicting external criteria over time.

Conclusions. The results support a dimensional representation of PDs that assesses both stable traits and dynamic processes.

INTRODUCTION

The DSM-IV (APA, 2000) criteria for personality disorders (Axis II PDs) conceptualize diagnoses as discrete categories. This model has demonstrated clinical utility (Cacciola *et al.* 1996; Tyrer *et al.* 2003; Casey *et al.* 2004; Gonzalez-Pinto *et al.* 2004) and predictive validity (Skodol *et al.* 1983; Gunderson *et al.* 1989; Smith *et al.* 1995; Histroke *et al.* 2003). However, limitations of this approach have been widely discussed and alternative dimensional models have been proposed (e.g. Widiger & Clark, 2000).

* Address for correspondence: Leslie C. Morey, Ph.D., Psychology Department, Texas A&M University, College Station, TX 77843-4235, USA.
(Email: lcm@psyc.tamu.edu)

Alternatives to DSM categories

One of the most widely studied alternatives is the Five-Factor Model (FFM), in which PDs are typically assumed to represent extreme or maladaptive variants of normal personality traits. Lower-order facets have been identified for the FFM to enhance its conceptual clarity and utility (Costa & McCrae, 1992), and applied to PDs (Lynam & Widiger, 2001). A number of studies have generally supported hypothesized associations with PDs for the FFM higher-order factors (Soldz *et al.* 1993; Blais, 1997; O'Connor & Dyce, 1998; Morey *et al.* 2000) and lower-order facets (Dyce & O'Connor, 1998; Trull *et al.* 2001; Morey *et al.* 2002; Warner *et al.* 2004). However, a meta-analysis of FFM–PD relationships as well as previous analyses from

our group suggest potential limitations to the FFM approach. In particular, the PDs are often characterized by a similar pattern involving high Neuroticism, low Conscientiousness, and low Agreeableness (Morey *et al.* 2000; Saulsman & Page, 2004). Thus, the FFM has demonstrated a stronger capacity to differentiate PDs from non-clinical controls than to differentiate PDs from each other.

A second dimensional model is represented by the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993), an instrument designed to assess personality characteristics in the normal and abnormal range. The SNAP was developed using an iterative approach that sought to identify the core dimensions underlying PD. The result was a measure with 12 trait scales indicative of personality pathology and three general temperament scales (Negative Temperament, Positive Temperament, Disinhibition). Although there has been less research on the relationships of SNAP scales to PDs than on the FFM, the available research has tended to support hypothesized associations (Reynolds & Clark, 2001; Morey *et al.* 2003; Stepp *et al.* 2005).

Predictive *versus* convergent validity

Many studies in this area treat DSM-IV PDs as a criterion standard (Livesley, 2001). The demonstration that the FFM and SNAP share content with a third model is a matter of convergent rather than predictive validity insofar as the DSM, FFM and SNAP represent three distinct efforts to conceptualize individuals characterized by similar features. In previous reports (Morey *et al.* 2002, 2003) we have noted convergences between these different approaches to personality description. However, common among theoretical systems of diagnostic acceptability (Robins & Guze, 1970; Kendell, 1975; Blashfield & Draguns, 1976) is the notion that predictive validity is more directly pertinent to the selection of diagnostic models than convergent validity, or, as Kendell (1975) noted, 'all diagnostic concepts stand or fall by the strength of the prognostic and therapeutic implications they embody' (p. 40). The demonstration that a diagnostic system is capable of reliably predicting a variety of etiological, concurrent and prognostic factors independent of and apart from other diagnostic symptoms is necessary to test hypotheses regarding relationships between models.

For example, Morey & Zanarini (2000) demonstrated that the FFM captures a sizable proportion of the variance in the DSM-IV borderline diagnosis with respect to antecedent, concurrent and predictive criteria in a large sample of in-patients. The DSM-IV diagnosis was largely unable to predict variance in 4-year outcome variables beyond that provided by the FFM but was found to be related to some important markers of diagnostic validity. Thus, it is important to move beyond questions regarding the convergence of all models to understanding the specific and relative utility of various personality models across a wide variety of important marker variables to accurately compare their underlying theories and make practical clinical determinations.

Purpose of the current study

The methodology used in the current study recognizes that no model has clearly demonstrated empirical superiority over the others in predicting clinical status and outcome in individuals with maladaptive personalities. Indicators of the models discussed above were administered to the Collaborative Longitudinal Study of Personality Disorders (CLPS) sample, which consists of individuals with PDs and major depressive disorder (MDD). This sample provides an opportunity to expand on previous validity studies of PD constructs that have been limited in terms of the number of comparison models and sample size, and that have focused primarily on single disorders. We selected the following model variants based on the theoretical issues described above: categorical and dimensional variants of DSM-IV Axis II, the constructs of the SNAP, and the FFM as represented by factor and facet level data. In contrast to a preliminary report published previously (Skodol *et al.* 2005), the present paper provides a comprehensive comparison of these personality models with respect to (1) prediction of antecedent, concurrent and predictive marker variables, (2) relative temporal stability, (3) degree of content overlap, and (4) incremental validity. Marker variables were chosen for their theoretical relationship to clinical outcome and PDs, and included childhood history, current and future functioning, treatment

utilization, suicidal behavior and Axis I psychopathology.

METHOD

Participants

Participants were 668 patients recruited from multiple clinical sites for the CLPS project (see Gunderson *et al.* 2000, for complete sample details). The study targeted four DSM-IV PDs, borderline (BPD), schizotypal (STPD), avoidant (AVPD) and obsessive-compulsive (OCPD), and a comparison group meeting criteria for MDD but with no PD. For this report, the original sample of 668 was reduced over time and across some measures because of attrition and incomplete data. Attrition was responsible for the reduction to 574 individuals at 2 years and 533 at 4 years who completed PD assessment interviews. In general, attriters tended to be slightly less impaired at baseline than those participants who were retained in the study (Warner *et al.* 2004).

Measures

Diagnostic Interview for DSM-IV Personality Disorders (*DIPD-IV*)

The DIPD-IV (Zanarini *et al.* 1996) is a semi-structured interview that assesses PD criteria that must be present over at least the previous 2 years to be counted towards the diagnosis. Adequate inter-rater reliability was found for all disorders diagnosed five times or more in a baseline subsample (Zanarini *et al.* 2000). The 10 DSM-IV Axis II diagnoses were included in regression models. Categorical diagnoses were determined by the cut-offs indicated in the DSM-IV, dimensional diagnoses by the number of clinically significant symptoms per diagnostic category.

Neuroticism Extraversion Openness Personality Inventory, Revised (*NEO-PI-R*)

The NEO-PI-R (Costa & McCrae, 1992) was designed to provide a comprehensive assessment of the five factors and 30 facets of the FFM. Internal consistency reliabilities for the five domains in this sample (Morey *et al.* 2002) ranged from 0·87 to 0·92. The scale scores from the five factors and 30 facets represent distinct models in this study because it has been suggested that an assessment of facets adds to the discriminative validity of the NEO-PI-R for PDs (Widiger *et al.* 1994).

Schedule for Nonadaptive and Adaptive Personality (*SNAP*)

The SNAP (Clark, 1993) is a 375-item self-report questionnaire designed to assess personality characteristics in both the normal and abnormal range. Internal consistency in our study sample was consistent with results described in the SNAP manual (Clark, 1993); medians of 0·89 for the higher-order temperament scales and 0·84 for the lower-order trait scales (Morey *et al.* 2003). Although the 12 'lower-order' traits are conceptualized as subcomponents of the three higher-order traits, higher-order factors are not a linear combination of lower-order scales and confirmatory factor analyses in our sample indicated that this hierarchical organization did not provide a good fit to the model (Hopwood & Morey, 2005). Thus, rather than adopting the SNAP 'three-factor' model, we used scale scores from all 15 dimensions to constitute the 'SNAP model' of personality. In this study, SNAP data gathered at baseline, 2-year and 3-year follow-up were examined; the SNAP was not administered at year 4.

Longitudinal Interval Follow-up Examination (*LIFE*)

The LIFE (Keller *et al.* 1987) is a structured interview that assesses functioning in interpersonal, recreational and occupational domains, the DSM-IV Global Assessment of Functioning (GAF), and several other outcome variables. Participants were assessed with the LIFE at baseline and after 24 and 48 months. The antecedent hospitalization variable was a retrospective estimate of the number of past hospitalizations; the 24- and 48-month hospitalization variables indicate the number of times participants had been hospitalized in the previous 2 years of follow-up. The antecedent medications variable represents a retrospective estimate of the number of psychiatric medications participants had ever been prescribed; the concurrent, 24- and 48-month medications variables represent the number of medications participants reported being prescribed at the time of the interview. Axis I disorders were rated

by clinicians as not present, in remission, present but subclinical, or clinically significant and used as dependent variables at 24 and 48 months. The 24- and 48-month suicide variables indicate the number of follow-up assessments in which participants reported at least one suicide attempt.

Social Adjustment Scale, Self-Report (SAS-SR)

The SAS-SR (Weissman & Bothwell, 1976) is a self-report instrument with adequate psychometric properties (Edwards *et al.* 1978) that yields estimates of interpersonal, occupational and recreational functioning. Median internal consistency reliability from CLPS baseline data for these composites was 0·70. Baseline, 24- and 36-month domain scores were used as dependent variables; the SAS-SR was not administered at year 4.

Childhood Experiences Questionnaire – Revised (CEQ-R)

The CEQ-R (Zanarini *et al.* 1989) is a clinician-rated interview that assesses a range of retrospectively reported childhood pathological and protective experiences, administered at baseline. Several composites theoretically related to PD etiology were included in the data analyses for this paper, including emotional denial; emotional, verbal, physical and sexual abuse; caretaker competence; positive relationships; and achievements. The median internal consistency reliability for the CEQ-R in the present sample was 0·83.

Structured Clinical Interview for DSM-IV Axis I (SCID)

The SCID (First *et al.* 1996) is a structured interview assessing DSM-IV Axis I disorders. For the current study, the Axis I variable was a sum score of clinically significant Axis I disorders as diagnosed based on the SCID. Axis I disorders counted towards the baseline Axis I criterion variable in this study included mood, anxiety, substance use and psychotic disorders.

Analyses

Data analyses proceeded in four stages representing the research questions posed above. To examine the diagnostic validity of the baseline diagnoses provided by the different classification models across a wide range of clinically relevant markers, the constructs of each model were related to antecedent, concurrent and predictive marker variables in a multiple regression framework. Temporal stability was also assessed, as this reflects an important assumption of personality traits and disorders. In addition, the abilities of each model to predict both itself and other models over time were calculated to determine associations between PD models. Finally, the incremental validity of each model, controlling for each other model, was calculated to indicate the unique relationship of different personality models to validity criteria.

Given this analytic strategy, it is important to note that the models compared in this study were characterized by different numbers of predictor variables (30 NEO-PI-R facets > 15 SNAP factors > 10 DSM-IV categories/dimensions > 5 NEO-PI-R factors). The tendency of non-significant predictor variables to artificially elevate multiple correlations in regression models comprising correlated predictor variables often result in over-fitting for models with more predictors. Although the adjusted R^2 is commonly used to correct for over-fitting, this method is purely algorithmic and thus penalizes models based on the number of variables and not on properties of the data themselves. To provide an empirical effect size estimate free from over-fitting, a Predicted Residual Sums of Squares (PRESS; Stevens, 2002) approach was used. In a PRESS analysis, a model is built with data from every participant except the one whose score is being predicted. This procedure occurs for each participant (i.e. the number of regression models equals the number of participants) and a cumulative effect size is estimated based on the observed residuals from this analysis for the entire sample. This provides a test of the ability of the model to generalize outside of the sample from which it is drawn, as the PRESS values are unaffected by the number of predictors in each model. However, an omnibus F test is not possible, and thus significance testing is described only for the full-sample models. Both regression- and PRESS-derived multiple correlations are reported for all analyses, although all substantive interpretations are based on PRESS coefficients. When appropriate (e.g. averaging correlations), Fisher's r to z transformation was used.

Table 1. *Multiple correlations of baseline personality models with external criteria*

Dependent variable	Categorical DSM-IV	Dimensional DSM-IV	SNAP	NEO-PI-R factors	NEO-PI-R facets
Antecedent ($n=668$)					
Emotional denial	**0·26** (0·17)	**0·30** (0·22)	**0·40** (0·33)	**0·26** (0·20)	**0·39** (0·17)
Emotional abuse	**0·24** (0·10)	**0·28** (0·20)	**0·39** (0·32)	**0·22** (0·14)	**0·36** (0·00)
Verbal abuse	**0·26** (0·14)	**0·28** (0·20)	**0·36** (0·26)	**0·20** (0·14)	**0·36** (0·00)
Physical abuse	**0·22** (0·00)	**0·24** (0·10)	**0·30** (0·20)	0·10 (0·00)	0·35 (0·00)
Sexual abuse	**0·32** (0·24)	**0·33** (0·24)	**0·35** (0·26)	0·14 (0·00)	**0·41** (0·17)
Caretaker competence	**0·28** (0·17)	**0·30** (0·22)	**0·30** (0·26)	**0·20** (0·10)	0·30 (0·00)
Positive relationships	0·20 (0·10)	**0·24** (0·20)	**0·28** (0·17)	**0·26** (0·20)	**0·41** (0·10)
Achievement	**0·35** (0·28)	**0·41** (0·36)	**0·36** (0·26)	**0·37** (0·33)	**0·48** (0·32)
Medications	**0·24** (0·17)	**0·24** (0·14)	**0·30** (0·17)	**0·20** (0·14)	**0·35** (0·00)
Hospitalizations	**0·26** (0·20)	**0·26** (0·17)	**0·30** (0·17)	0·10 (0·00)	0·28 (0·00)
Baseline ($n=668$)					
GAF	**0·47** (0·44)	**0·57** (0·55)	**0·54** (0·50)	**0·41** (0·39)	**0·51** (0·40)
LIFE interpersonal	**0·44** (0·39)	**0·51** (0·49)	**0·44** (0·39)	**0·36** (0·32)	**0·45** (0·30)
LIFE work	**0·37** (0·22)	**0·42** (0·28)	**0·42** (0·20)	**0·26** (0·22)	**0·35** (0·00)
LIFE recreation	**0·33** (0·26)	**0·40** (0·35)	**0·46** (0·40)	**0·35** (0·32)	**0·47** (0·33)
SAS interpersonal	**0·35** (0·30)	**0·40** (0·36)	**0·54** (0·50)	**0·45** (0·42)	**0·52** (0·41)
SAS work	**0·24** (0·14)	**0·28** (0·22)	**0·36** (0·28)	**0·30** (0·26)	**0·40** (0·24)
SAS recreation	**0·41** (0·37)	**0·52** (0·49)	**0·59** (0·57)	**0·49** (0·47)	**0·57** (0·49)
Axis I	**0·36** (0·30)	**0·46** (0·42)	**0·47** (0·42)	**0·40** (0·37)	**0·39** (0·24)
Medications	**0·36** (0·30)	**0·39** (0·33)	**0·41** (0·36)	**0·26** (0·24)	**0·47** (0·33)
24 months ($n=574$)					
GAF	**0·44** (0·40)	**0·51** (0·48)	**0·47** (0·42)	**0·40** (0·36)	**0·48** (0·33)
LIFE interpersonal	**0·40** (0·33)	**0·49** (0·45)	**0·37** (0·30)	**0·33** (0·30)	**0·40** (0·17)
LIFE work	0·22 (0·00)	**0·24** (0·14)	**0·30** (0·17)	**0·32** (0·26)	**0·42** (0·22)
LIFE recreation	**0·24** (0·10)	**0·30** (0·22)	**0·33** (0·24)	**0·35** (0·32)	**0·42** (0·22)
SAS interpersonal	**0·32** (0·22)	**0·40** (0·33)	**0·45** (0·36)	**0·39** (0·36)	**0·45** (0·17)
SAS work	**0·26** (0·24)	**0·30** (0·20)	**0·41** (0·32)	**0·33** (0·28)	**0·47** (0·20)
SAS recreation	**0·40** (0·33)	**0·50** (0·45)	**0·57** (0·52)	**0·46** (0·42)	**0·52** (0·35)
Suicide attempts	**0·37** (0·36)	**0·41** (0·40)	**0·44** (0·42)	**0·26** (0·20)	**0·39** (0·00)
Axis I	**0·41** (0·35)	**0·47** (0·42)	**0·48** (0·42)	**0·42** (0·39)	**0·47** (0·28)
Hospitalizations	**0·28** (0·14)	**0·26** (0·17)	**0·26** (0·10)	0·14 (0·10)	0·35 (0·00)
Medications	**0·32** (0·20)	**0·36** (0·24)	**0·40** (0·26)	**0·28** (0·22)	**0·41** (0·10)
48 months ($n=533$)					
GAF	**0·33** (0·28)	**0·44** (0·39)	**0·46** (0·40)	**0·42** (0·40)	**0·51** (0·37)
LIFE interpersonal	**0·36** (0·28)	**0·42** (0·37)	**0·28** (0·35)	**0·36** (0·32)	**0·44** (0·22)
LIFE work	**0·24** (0·14)	**0·28** (0·20)	**0·28** (0·28)	**0·36** (0·32)	**0·47** (0·28)
LIFE recreation	**0·24** (0·10)	**0·28** (0·17)	**0·37** (0·28)	**0·36** (0·32)	**0·44** (0·22)
SAS interpersonal[a]	**0·26** (0·10)	**0·35** (0·24)	**0·46** (0·37)	**0·35** (0·28)	**0·46** (0·14)
SAS work[a]	0·24 (0·10)	**0·30** (0·17)	**0·42** (0·33)	**0·28** (0·22)	**0·42** (0·00)
SAS recreation[a]	**0·40** (0·33)	**0·47** (0·41)	**0·53** (0·47)	**0·50** (0·47)	**0·55** (0·37)
Suicide attempts	**0·24** (0·20)	**0·26** (0·17)	**0·28** (0·10)	0·10 (0·10)	0·26 (0·00)
Axis I	**0·35** (0·26)	**0·39** (0·33)	**0·45** (0·39)	**0·40** (0·37)	**0·49** (0·32)
Hospitalizations	0·20 (0·10)	**0·24** (0·14)	**0·26** (0·14)	0·20 (0·10)	0·32 (0·17)
Medications	**0·30** (0·14)	**0·33** (0·20)	**0·37** (0·20)	**0·22** (0·14)	**0·39** (0·00)

SNAP, Schedule for Nonadaptive and Adaptive Personality; NEO-PI-R, Neuroticism Extraversion Openness Personality Inventory, Revised; GAF, Global Assessment of Functioning; LIFE, Longitudinal Interval Follow-up Examination; SAS, Social Adjustment Scale.

[a] SAS variables (Interpersonal, Work and Recreation Composites) measured at 36 months.

Significance for multiple regression ($p<0·001$) indicated by bold, Predicted Residual Sums of Squares (PRESS) r in parentheses.

RESULTS

The first set of analyses involved the use of five models of personality pathology to predict relevant criterion variables. The results of these analyses are reported as both multiple correlations and the square root of PRESS r^2 values (see Table 1). A type I error rate of 0·001 was selected to adjust for the probability of spuriously significant values in multiple analyses (there were 42 criteria; a Bonferroni adjustment of a type I error rate of 0·05 suggests $p=0·05/42=0·001$).

Model validity

There was substantial support for the validity of all five models in predicting a wide range of criteria, with 194 of 210 regression models achieving significance. Independent sign tests

indicated that the SNAP produced larger correlations with significantly more criteria ($p<0{\cdot}001$) than the other models. The dimensional DSM-IV diagnosis obtained a larger PRESS value than the categorical diagnosis in predicting 36 of 42 dependent variables ($p<0{\cdot}001$). The FFM factors had a larger PRESS value than the facets in predicting 33 variables, whereas the facets had a larger PRESS value in five, a statistically significant difference (sign test, $p<0{\cdot}001$). Finally, the dimensionalized DSM-IV obtained a larger PRESS value for twice as many validity criteria as the FFM factors; this difference approached statistical significance ($p<0{\cdot}10$).

The models appeared to differ with respect to the influence of time on criterion validity. Specifically, the average difference between the comparable 48-month and baseline PRESS values was larger for the DSM-IV (average decrease in multiple correlation = 0·11) than for the SNAP (0·06) or FFM (0·04) values. This difference indicates that the SNAP and FFM are able to sustain their predictive validity over time to a greater degree than the DSM-IV.

The median PRESS values for FFM domains, dimensionalized DSM-IV and SNAP models, across all outcome criteria for baseline, 24-month and 48-month time points are depicted in Fig. 1. The SNAP predicts more variability in baseline criteria than the DSM-IV, which in turn predicts more variability in baseline criteria than the FFM. At 24 months, this gap decreases and all models are essentially equal. However, a dramatic decrease in predictive validity is observed for the DSM-IV at 48 months, whereas the SNAP and FFM retain their validity. These data suggest that the strengths of the DSM-IV (concurrent validity) and the FFM (predictive validity) are both incorporated to some degree in the SNAP. To further examine this interpretation, a hybrid model comprising 10 DSM-IV dimensional scores and five FFM factor scores was tested across all criterion variables. This 15-variable hybrid model demonstrated larger validity coefficients than the 15-variable SNAP model (0·30 for FFM/DSM-IV *v.* 0·26 for SNAP in predicting antecedent variables, 0·48 *v.* 0·40 at baseline, 0·40 *v.* 0·32 at 24 months, and 0·40 *v.* 0·33 at 48 months). These results suggest that it may be this hybrid of static/normative and dynamic/maladaptive variables,

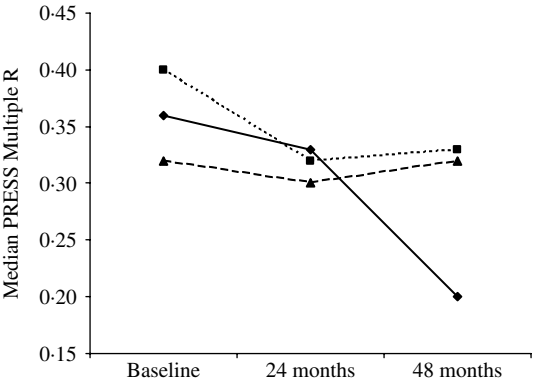

FIG. 1. Median Predicted Residual Sums of Squares (PRESS) multiple correlation coefficients across all validating variables for three baseline models over time. DSM, dimensionalized DSM-IV Axis II criteria (–◆–); SNAP, Schedule for Nonadaptive and Adaptive Personality (···■···); FFM, Five-Factor Model (--▲--).

rather than the particular constructs in the model, that leads to optimal validity in predicting clinically relevant criteria.

Model stability

Baseline to 24-month and baseline to 48-month (36-month for SNAP) correlations were computed for DSM-IV dimensional diagnoses, FFM factors and SNAP traits (Table 2) to assess the temporal stability of each model. All of the models and their subscales/diagnoses demonstrated large and statistically significant ($p<0{\cdot}01$) correlations over time. The baseline to 4-year correlations on the FFM facets were larger than the correlations on the DSM-IV Axis II diagnoses over the same time period [$t(39)=6{\cdot}23$, $p<0{\cdot}001$]. SNAP data suggested that, like the FFM, it is tapping enduring personality traits.

Model convergence

Table 3 depicts the ability of the baseline dimensional and categorical DSM-IV diagnoses, SNAP traits and FFM factors and facets to predict the number of Axis II symptoms, SNAP temperament scales and FFM factors at baseline, 24 and 48 months (36 months for SNAP, which was not administered at 48 months). All models demonstrated impressive validity in predicting each other, which supports the assumption that they are accounting for overlapping variance.

Table 2. *Personality model stability*

	Baseline to 24 months	Baseline to 48 months
DSM dimensional diagnosis		
Borderline	0·61	0·48
Schizotypal	0·68	0·51
Avoidant	0·62	0·49
Obsessive-compulsive	0·52	0·43
Mean (all 10)[a]	0·59	0·47
NEO-PI-R		
Neuroticism	0·68	0·66
Extraversion	0·71	0·69
Openness	0·78	0·79
Agreeableness	0·77	0·74
Conscientiousness	0·75	0·74
Mean (facets)[a]	0·67	0·63
SNAP[b]		
Negative temperament	0·61	0·62
Mistrust	0·75	0·74
Manipulativeness	0·69	0·59
Aggression	0·72	0·70
Self-harm	0·69	0·69
Eccentric perceptions	0·65	0·63
Dependency	0·57	0·56
Positive temperament	0·69	0·70
Exhibitionism	0·76	0·71
Entitlement	0·68	0·63
Detachment	0·68	0·64
Disinhibition	0·67	0·73
Impulsivity	0·70	0·67
Propriety	0·68	0·67
Workaholism	0·67	0·66

NEO-PI-R, Neuroticism Extraversion Openness Personality Inventory, Revised; SNAP, Schedule for Nonadaptive and Adaptive Personality.

[a] Scale correlations were transformed to *z*-scores, added, and transformed back to correlations.

[b] Second column indicates 0–36-month correlations.

Incremental validity

Given that models demonstrated validity in predicting a wide range of external criteria over time (Table 1) and were substantially related to one another (Table 3), we investigated the extent to which each model increments the others in explaining validity markers. To focus this investigation, specific models were selected based upon results depicted in Table 1. The dimensional model of the DSM-IV diagnoses was examined because it consistently outperformed the categorical model; the FFM factor model was selected over the facet model because the facets did not appear to provide incremental validity to the factors in terms of PRESS values. A series of steps were taken to remove the artificial increase in correlations associated with larger numbers of predictors in regression analyses. First, PRESS predicted scores were computed for participants for each of the three models, predicting each validity criterion, thus allowing each model to be represented by one independent variable that reflects its estimate of each participant's criterion score. Second, two of these estimates were entered hierarchically into a regression model predicting each validity criterion, and the change in the multiple correlation resulting from adding the estimate from the second classification model was calculated. This multiple correlation represents the part-correlation of each model with the criterion after removing the variance in the criterion explained by the other model and indicates the incremental validity added by the first model in estimating the specified criterion over and above the estimate provided by the second model.

The results of these analyses are reported in Table 4. Statistically significant part-correlations were observed between the SNAP and most of the criteria after controlling for the contribution of the DSM-IV. However, less than half of the part-correlations of the DSM-IV with criteria after controlling the SNAP are statistically significant. The results also indicated that the SNAP incremented the information provided by the FFM to a greater extent than was true of the converse. It seems that the DSM-IV tends to add significant incremental validity to the FFM in the prediction of antecedent and baseline variables (whereas the opposite is generally not the case), and that each model incremented the other at 24 and 48 months. In addition, whereas the part-correlations appear to decrease for the DSM-IV as a function of time across criteria, regardless of the model being controlled for, the opposite is true for the FFM.

DISCUSSION

This study sought to compare the relative utility of three general models (and variants of two of these models) in predicting a variety of theoretically and clinically relevant variables related to PD. All models (categorical and dimensional DSM-IV, SNAP and factor and facet level FFM) demonstrated significant stability and were able to predict a wide variety of theoretically relevant variables over time. However, the results also indicated that these models have different strengths and weaknesses.

Table 3. *Correlations of baseline models predicting model variables over time*

Dependent variable	Categorical DSM-IV	Dimensional DSM-IV	SNAP	NEO-PI-R factors	NEO-PI-R facets
Baseline					
Axis II disorders	—	—	**0·67** (0·65)	**0·58** (0·57)	**0·65** (0·57)
Positive temperament	**0·45** (0·40)	**0·53** (0·50)	—	**0·77** (0·76)	**0·82** (0·80)
Negative temperament	**0·46** (0·42)	**0·54** (0·52)	—	**0·78** (0·77)	**0·84** (0·81)
Disinhibition	**0·44** (0·40)	**0·55** (0·52)	—	**0·73** (0·72)	**0·80** (0·76)
Neuroticism	**0·53** (0·51)	**0·62** (0·60)	**0·85** (0·84)	—	—
Extraversion	**0·48** (0·45)	**0·60** (0·57)	**0·85** (0·84)	—	—
Openness to experience	0·20 (0·00)	**0·28** (0·22)	**0·52** (0·48)	—	—
Agreeableness	**0·46** (0·42)	**0·59** (0·57)	**0·77** (0·75)	—	—
Conscientiousness	**0·37** (0·32)	**0·45** (0·40)	**0·82** (0·81)	—	—
24 months					
Axis II disorders	**0·39** (0·35)	**0·51** (0·48)	**0·34** (0·29)	**0·28** (0·25)	**0·35** (0·21)
Positive temperament	**0·33** (0·24)	**0·39** (0·33)	**0·71** (0·68)	**0·56** (0·54)	**0·62** (0·50)
Negative temperament	**0·37** (0·32)	**0·45** (0·40)	**0·65** (0·62)	**0·56** (0·54)	**0·60** (0·47)
Disinhibition	**0·39** (0·32)	**0·46** (0·41)	**0·75** (0·69)	**0·60** (0·57)	**0·69** (0·59)
Neuroticism	**0·42** (0·36)	**0·47** (0·42)	**0·65** (0·60)	**0·69** (0·67)	**0·70** (0·61)
Extraversion	**0·39** (0·32)	**0·47** (0·42)	**0·67** (0·63)	**0·71** (0·70)	**0·73** (0·65)
Openness to experience	**0·20** (0·00)	**0·24** (0·00)	**0·49** (0·41)	**0·75** (0·74)	**0·79** (0·73)
Agreeableness	**0·45** (0·40)	**0·57** (0·53)	**0·69** (0·65)	**0·77** (0·76)	**0·79** (0·73)
Conscientiousness	**0·36** (0·26)	**0·39** (0·30)	**0·64** (0·60)	**0·76** (0·75)	**0·78** (0·71)
48 months					
Axis II disorders	**0·25** (0·18)	**0·35** (0·29)	**0·33** (0·25)	**0·23** (0·19)	**0·33** (0·12)
Positive temperament[a]	**0·36** (0·28)	**0·40** (0·33)	**0·71** (0·68)	**0·58** (0·56)	**0·66** (0·56)
Negative temperament[a]	**0·37** (0·30)	**0·42** (0·36)	**0·65** (0·61)	**0·59** (0·57)	**0·63** (0·52)
Disinhibition[a]	**0·35** (0·24)	**0·52** (0·42)	**0·75** (0·72)	**0·62** (0·59)	**0·71** (0·62)
Neuroticism	**0·41** (0·33)	**0·45** (0·39)	**0·63** (0·59)	**0·67** (0·66)	**0·69** (0·60)
Extraversion	**0·37** (0·28)	**0·46** (0·40)	**0·66** (0·62)	**0·69** (0·67)	**0·71** (0·63)
Openness to experience	0·20 (0·00)	0·22 (0·00)	**0·48** (0·40)	**0·77** (0·75)	**0·78** (0·72)
Agreeableness	**0·37** (0·30)	**0·52** (0·48)	**0·62** (0·58)	**0·73** (0·71)	**0·77** (0·71)
Conscientiousness	**0·33** (0·24)	**0·35** (0·26)	**0·64** (0·60)	**0·75** (0·73)	**0·78** (0·72)

NEO-PI-R, Neuroticism Extraversion Openness Personality Inventory, Revised; SNAP, Schedule for Nonadaptive and Adaptive Personality; PRESS, Predicted Residual Sums of Squares.

[a] SNAP variables (positive temperament, negative temperament, disinhibition) measured at 36 months. Significance for multiple regression ($p < 0{\cdot}001$) indicated by bold, PRESS r in parentheses.

All of the considered dimensional models, including a dimensional conceptualization of the DSM-IV, consistently outperformed the commonly used categorical diagnosis. In the book *A Research Agenda for DSM-V* (Kupfer *et al.* 2002), the question of whether mental illness should be represented by dimensions or categories is identified as one of the seven basic nomenclature issues needing clarification. Rounsaville *et al.* (2002) state 'there is a clear need for dimensional models to be developed and their utility compared with that of existing typologies in one or more limited fields, such as personality' (p. 13). They note that it is an open question whether dimensional models will be successful in clinical settings. Evidence exists that categorical personality diagnosis is considerably less reliable than dimensional diagnosis with respect to both interdiagnostic agreement (Heumann & Morey, 1989) and temporal stability (Grilo *et al.* 2004) and that, as seen here, categorical diagnosis of PD also appears to be considerably less valid. The data described here are consistent with the common opinion that moving the DSM-V into a dimensional system for classifying PDs seems timely.

The data concerning the relative validity of the alternative dimensional models examined in this study are less clear-cut, as each appeared to have significant validity. The SNAP model demonstrated greater capacity, relative to the FFM and the DSM-IV, to predict external markers of construct validity. This trend was evident across a variety of validity criteria from different domains (e.g. etiologic, functional, symptomatic) and from different measurement methods (e.g. interview *versus* self-report validity indicators). The SNAP was also able to significantly increment the other models in predicting a wide range of criteria, while the opposite was, for the most part, not true. The

Table 4. *Part-correlations of PRESS-derived predicted scores with external criteria*

	DSM-IV		SNAP		NEO-PI-R	
	SNAP	NEO-PI-R	DSM	NEO-PI-R	DSM	SNAP
Antecedent						
Emotional denial	0·07	**0·14**	**0·25**	**0·26**	0·10	0·00
Emotional abuse	0·05	**0·15**	**0·25**	**0·25**	0·07	0·01
Verbal abuse	0·07	**0·15**	**0·20**	**0·22**	0·08	0·00
Physical abuse	0·06	**0·13**	**0·18**	**0·19**	0·04	0·06
Sexual abuse	**0·14**	**0·23**	**0·19**	**0·24**	0·00	0·00
Positive experiences	0·09	0·04	**0·13**	0·10	**0·19**	**0·13**
Caretaker competence	**0·18**	**0·16**	**0·13**	0·08	0·06	0·10
Achievements	**0·26**	**0·25**	0·12	0·06	**0·17**	**0·20**
Past medications	0·08	**0·15**	**0·15**	**0·22**	0·10	0·00
Hospitalizations	0·13	**0·23**	**0·13**	**0·25**	0·12	0·09
Baseline						
GAF	**0·29**	**0·42**	**0·22**	**0·38**	**0·14**	0·05
LIFE interpersonal	**0·31**	**0·34**	**0·13**	**0·24**	0·07	0·10
LIFE work	**0·20**	**0·21**	0·10	**0·13**	0·11	0·09
LIFE recreation	**0·15**	**0·19**	**0·26**	**0·27**	0·10	0·00
SAS interpersonal	0·00	0·12	**0·36**	**0·23**	**0·31**	0·12
SAS work	0·07	**0·16**	**0·20**	**0·15**	**0·17**	0·09
SAS recreation	**0·15**	**0·24**	**0·33**	**0·29**	**0·24**	0·11
No. Axis I disorders	**0·19**	**0·24**	**0·20**	**0·22**	**0·19**	0·11
Current medications	**0·15**	**0·29**	**0·20**	**0·32**	0·08	0·00
24 months						
GAF	**0·26**	**0·34**	**0·18**	**0·33**	0·13	0·07
LIFE interpersonal	**0·33**	**0·32**	0·10	**0·18**	0·09	0·11
LIFE work	0·06	0·13	**0·17**	0·13	**0·18**	0·09
LIFE recreational	0·10	0·10	**0·16**	0·10	**0·23**	**0·16**
SAS interpersonal	0·12	**0·21**	**0·22**	**0·21**	**0·18**	0·10
SAS work	0·06	0·13	**0·25**	**0·25**	**0·20**	0·05
SAS recreational	**0·17**	**0·26**	**0·31**	**0·31**	**0·23**	0·08
No. Axis I disorders	**0·18**	**0·27**	**0·20**	**0·25**	**0·16**	0·09
Suicide attempts	**0·17**	**0·29**	**0·24**	**0·33**	0·07	0·00
Hospitalizations	0·13	**0·17**	0·11	**0·19**	0·02	0·02
Medications	**0·14**	**0·20**	**0·19**	**0·24**	0·13	0·03
48 months						
GAF	**0·19**	**0·26**	**0·20**	**0·25**	**0·23**	0·09
LIFE interpersonal	**0·20**	**0·20**	**0·18**	**0·18**	**0·16**	0·11
LIFE work	0·04	0·08	**0·23**	0·11	**0·26**	0·14
LIFE recreational	0·04	0·02	**0·23**	0·11	**0·27**	0·13
SAS interpersonal	0·03	**0·15**	**0·29**	**0·30**	**0·17**	0·04
SAS work	0·00	**0·16**	**0·29**	**0·29**	**0·17**	0·03
SAS recreational	**0·16**	**0·20**	**0·28**	**0·23**	**0·27**	**0·16**
No. Axis I disorders	0·11	**0·16**	**0·23**	**0·18**	**0·21**	0·09
Suicide attempts	0·12	**0·24**	0·09	0·13	0·04	0·03
Hospitalizations	0·10	0·14	0·11	**0·18**	0·10	0·07
Medications	0·11	**0·19**	**0·17**	0·04	0·09	0·00

PRESS, Predicted Residual Sums of Squares; NEO-PI-R, Neuroticism Extraversion Openness Personality Inventory, Revised; SNAP, Schedule for Nonadaptive and Adaptive Personality; GAF, Global Assessment of Functioning; LIFE, Longitudinal Interval Follow-up Examination; SAS, Social Adjustment Scale.

Significance ($p < 0{\cdot}001$) indicated by bold.

relative success of the SNAP constructs over those of the DSM-IV and FFM appeared to be related to its ability to combine the relative strengths of these two alternatives. The DSM-IV diagnostic model was strongest in its ability to capture the pattern of functional impairment at the time that this assessment was made, consistent with our previous analyses (Skodol *et al*. 2005). However, this ability gradually diminished over time, indicating that the DSM-IV information is particularly pertinent to current mental, emotional, environmental and behavioral states, but these phenomena are less stable over time and hence less able to permit longer-term predictions. By contrast, the FFM traits provide less information about current

behaviors and functional patterns, but the information they do provide reflects more stable and enduring aspects of personality that continue to be predictive well into the future. These findings suggest that PDs are a hybrid of two elements (McGlashan *et al.* 2005): (*a*) stable personality traits that may have normal variants but that in PDs may be pathologically exaggerated, and (*b*) less stable dysfunctional or maladaptive characteristics that may reflect attempts at adapting to, coping with, or compensating for these traits as manifest in a particular life situation or context. This hypothesis was further supported by supplemental analyses that demonstrated that a hybrid combination of FFM and DSM-IV constructs compared favorably to those of the SNAP.

Recently, integrative models involving the FFM, SNAP and other personality models have been proposed to clarify the broadest levels of the personality trait hierarchy (Markon *et al.* 2005; Trull & Durrett, 2005). Based on the consistent finding of significant overlap across theoretical and measurement models, it has been argued that most of the important variation in personality is contained within the trait hierarchy. This interpretation seems justified by the adequacy of structural models that integrate a variety of instruments (e.g. Markon *et al.* 2005). However, an important question involves whether additional predictive variance lies at lower levels of the trait hierarchy or is better described as being conceptually independent of that hierarchy. The present study suggested that lower-order features in a normative trait hierarchy (the FFM facets) failed to improve upon the higher-order factors in predicting validating variables. Conversely, adding the more maladaptively focused DSM-IV PD constructs to the higher-order FFM model did make it comparable, if not superior, to the SNAP in terms of validity. These results, combined with earlier analyses suggesting that the scales of the SNAP do not seem to support the hypothesized hierarchical structure (Hopwood & Morey, 2005), suggest that the relationship between normative trait dimensions and maladaptive personality features is not entirely one of hierarchical superordinacy.

Several alternative explanations of study findings should be considered. It is possible that there is an interaction between the PDs studied and the models being tested, such that some models may be particularly effective at capturing aspects of the four PDs targeted in this study, which are over-represented in this sample by design. Instrument method effects might have favored the DSM-IV approach, as most of the criterion variables were assessed by interview, as was the DSM-IV but not the FFM or SNAP, which were assessed by self-report. Conversely, most of the sample were in ongoing treatment during data collection, particularly during the early stages of the study (Bender *et al.* 2001), and this treatment presumably targeted behaviors and traits represented by the DSM PD diagnoses, perhaps leading to greater change in these behaviors and rendering them less able to predict future status. Measurement issues may have also negatively affected the potential of the FFM because NEO-PI-R items were selected and written to assess non-clinical populations; they may not make useful discriminations at the extremes of each of the five factors (Haigler & Widiger, 2001). It may be that an instrument that assesses the more extreme range of FFM lower-order traits (e.g. Trull & Widiger, 1997) would predict much of what is captured by the SNAP and DSM-IV but missed by the facets and factors of the NEO-PI-R. Finally, although efforts were made to sample a variety of important clinical markers, the universe of variables a clinician may want to predict is not fully represented in the current report.

At a broad level, two results of this study are particularly salient for the classification of PD. First, the results of this study provide consistent demonstration that dimensional characterizations of PD provide information beyond that given by the commonly used categorical diagnosis. Second, models of personality pathology that represent stable trait dispositions as well as more dynamic, maladaptive behavioral manifestations of personality pathology are most clinically informative. Both of these results have important implications for DSM-V. The least radical change from the current format that would accomplish these goals would involve (*a*) dimensionalizing current PD categories in some manner, and (*b*) adding a formal assessment of normative personality traits identified in factor analytic work. It is likely that weaving in an assessment of higher-order personality traits

would be informative for those that are currently categorized as Axis II as well as Axis I disorders (Hopwood *et al*. in press). An ultimate goal should involve the elaboration of a truly integrative representation of PD that clarifies (*a*) the entire spectrum of the personality trait hierarchy at both higher and lower orders, and (*b*) clinically relevant phenomena that are not subsumed within that hierarchy, perhaps including disease processes or environmental, situational or developmental factors.

ACKNOWLEDGMENTS

This work was supported by NIMH grants MH 50837, 50838, 50839, 50840, 50850 and MH01654 (T.H.M.). This publication has been reviewed and approved by the Publications Committee of the Collaborative Longitudinal Personality Disorders Study.

DECLARATION OF INTEREST

None.

REFERENCES

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn, revised). American Psychiatric Association: Washington, DC.

Bender, D. S., Dolan, R. T., Skodol, A. E., Sanislow, C. A., Dyck, I. R., McGlashan, T. H., Shea, M. T., Zanarini, M. C., Oldham, J. M. & Gunderson, J. G. (2001). Treatment utilization by patients with personality disorders. *American Journal of Psychiatry* **158**, 295–302.

Blais, M. A. (1997). Clinician rating of the five-factor model of personality and the DSM-IV personality disorders. *Journal of Nervous and Mental Disease* **185**, 388–393.

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

Cacciola, J. S., Rutherford, M. J., Alterman, A. I., McKay, J. R. & Snider, E. C. (1996). Personality disorders and treatment outcome in methadone maintenance patients. *Journal of Nervous and Mental Disease* **184**, 234–239.

Casey, P, Birbeck, C., McDonagh, C., Horgan, A., Dirwick, C., Delgard, O., Lethinin, V., Ayusuo-Mateos, J. L., Dunn, G., Page, H., Wilkinson, C., Wilkinson, G. & Vasquez-Barquero, J. L. (2004). Personality disorder, depression, and functioning: results from the ODIN study. *Journal of Affective Disorders* **82**, 277–283.

Clark, L. A. (1993). *Manual for the Schedule of Nonadaptive and Adaptive Personality*. University of Minnesota Press: Minneapolis, MN.

Costa, P. T. & McRae, R. R. (1992). *Professional Manual: Revised NEO Personality Inventory* (*NEO-PI-R*) *and NEO Five-Factor Inventory* (*NEO-FFI*). Psychological Assessment Resources: Odessa, FL.

Dyce, J. A. & O'Connor, B. P. (1998). Personality disorders and the five-factor model: a test of facet-level predictions. *Journal of Personality Disorders* **12**, 31–45.

Edwards, D. W., Yarvis, R. M., Mueller, D. P., Zingale, H. C. & Wagman, W. J. (1978). Test-taking and the stability of adjustment scales: can we assess patient deterioration? *Evaluation Quarterly* **2**, 275–292.

First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders, Patient Edition* (*SCID-P*). New York State Psychiatric Institute, Biometrics Research: New York.

Gonzalez-Pinto, A., Immaculada, F., Rodriguez, C., de Corres Blanco, F. & Sonsoles, E. (2004). Purging behavior and comorbidity as predictors of quality of life in anorexia nervosa. *International Journal of Eating Disorders* **36**, 445–450.

Grilo, C. M., Shea, M. T., Sanislow, C. A., Skodol, A. E., Stout, R. L., Gunderson, J. G., Yen, S., Bender, D. S., Pagano, M. E., Zanarini, M. C., Morey, L. C. & McGlashan, T. H. (2004). Two-year stability and change in schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology* **72**, 767–775.

Gunderson, J. G., Frank, A. F., Ronningstam, E. F., Wachter, S., Lynch, V. J. & Wolf, P. J. (1989). Early discontinuance of borderline patients from psychotherapy. *Journal of Nervous and Mental Disease* **177**, 38–42.

Gunderson, J. G., Shea, M. T., Skodol, A. E., McGlashan, T. H., Morey, L. C., Stout, R. L., Zanarini, M. C., Grilos, C. M., Oldham, J. M. & Keller, M. B. (2000). The Collaborative Longitudinal Personality Disorders Study: development, aims, design, and sample characteristics. *Journal of Personality Disorders* **14**, 300–315.

Haigler, E. D. & Widiger, T. A. (2001). Experimental manipulation of NEO-PI-R items. *Journal of Personality Assessment* **77**, 339–358.

Heumann, K. A. & Morey, L. C. (1990). Reliability of categorical and dimensional judgments of personality disorder. *American Journal of Psychiatry* **147**, 498–500.

Histroke, U., Langstrom, N., Ottosson, H. & Grann, M. (2003). Self-reported personality traits and disorders (DSM-IV) and risk of criminal recidivism: a prospective study. *Journal of Personality Disorders* **17**, 293–305.

Hopwood, C. J. & Morey, L. C. (2005). Predictive validity of the SNAP in a large, personality-disordered sample. Society for Personality Assessment Annual Conference, Chicago, IL.

Hopwood, C. J., Morey, L. C., Shez, M. T., McGlashan, T. H., Stanislow, C. A., Grilo, C. M., Gunderson, J. G., Zanarini, M. C. & Skodol, A. E. (in press). Personality traits predict current and future functioning comparably for individuals with major depressive and personality disorders. *Journal of Nervous and Mental Disease*.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P. & Andreason, N. C. (1987). The longitudinal interval follow-up evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

Kendell, R. E. (1975). *The Role of Diagnosis in Psychiatry*. Blackwell Scientific Publications: Oxford.

Kupfer, D. J., First, M. B. & Regier, D. A. (2002). *A Research Agenda for DSM-V*. American Psychiatric Press: Washington, DC.

Livesley, W. J. (2001). Commentary on reconceptualizing personality disorder categories using trait dimensions. *Journal of Personality* **69**, 277–286.

Lynam, D. R. & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: an expert consensus approach. *Journal of Abnormal Psychology* **110**, 401–412.

Markon, K. E., Krueger, R. F. & Watson, D. (2005). Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *Journal of Personality and Social Psychology* **88**, 139–157.

McGlashan, T. H., Grilo, C. M., Sanislow, C. A., Ralevski, E., Morey, L. C., Gunderson, J. G., Skodol, A. E., Shea, M. T., Zanarini, M. C., Bender, D., Stout, R. L., Yen, S. & Pagano, M. (2005). Two-year prevalence and stability of individual DSM-IV

criteria for schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *American Journal of Psychiatry* **162**, 883–889.

Morey, L. C., Gunderson, J. G., Quigley, B. A. & Lyons, M. (2000). Dimensions and categories: The 'big five' factors and the DSM personality disorders. *Assessment* **7**, 203–216.

Morey, L. C., Gunderson, J. G., Quigley, B. D., Shea, M. T., Skodol, A. E., McGlashan, T. H., Stout, R. L. & Zanarini, M. C. (2002). The representation of borderline, avoidant, obsessive-compulsive, and schizotypal personality disorders by the five-factor model. *Journal of Personality Disorders* **16**, 215–234.

Morey, L. C., Warner, M. B., Shea, M. T., Gunderson, J. B., Sanislow, C. A., Grilo, C., Skodol, A. E. & McGlashan, T. H. (2003). The representation of four personality disorders by the Schedule for Nonadaptive and Adaptive Personality dimensional model of personality. *Psychological Assessment* **15**, 326–332.

Morey, L. C. & Zanarini, M. C. (2000). Borderline personality: traits and disorder. *Journal of Abnormal Psychology* **109**, 733–737.

O'Connor, B. P. & Dyce, J. A. (1998). A test of models of personality disorder configuration. *Journal of Abnormal Psychology* **107**, 3–16.

Reynolds, S. K. & Clark, L. A. (2001). Predicting dimensions of personality disorder from domains and facets of the five-factor model. *Journal of Personality* **69**, 199–222.

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

Rounsaville, B. J., Alarcon, R. D., Andrews, G., Jackson, J. S., Kendell, R. E. & Kendler, K. (2002). Basic nomenclature issues for DSM-V. In *A Research Agenda for DSM-V* (ed. D. J. Kupfer, M. B. First and D. A. Regier), pp. 1–29. American Psychiatric Association: Washington, DC.

Saulsman, L. M. & Page, A. C. (2004). The five-factor model and personality disorder empirical literature: a meta-analytic review. *Clinical Psychology Review* **23**, 1055–1085.

Skodol, A. E., Buckley, P. & Charles, E. (1983). Is there a characteristic pattern to the treatment history of clinic outpatients with borderline personality? *Journal of Nervous and Mental Disease* **171**, 405–410.

Skodol, A. E., Oldham, J. E., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C., Shea, M. T., Zanarini, M. C., Sanislow, C. A., Grilo, C. M., McGlashan, T. H. & Gunderson, J. G. (2005). Dimensional representations of DSM-IV personality disorders: relationships to functional impairment. *American Journal of Psychiatry* **162**, 1919–1925.

Smith, T. E., Koenigsberg, H. E., Yeomans, F. E. & Clarkin, J. F. (1995). Predictors of dropout in psychodynamic psychotherapy of borderline personality disorder. *Journal of Psychotherapy Practice and Research* **4**, 205–213.

Soldz, S., Budman, S., Demby, A. & Merry, J. (1993). Representation of personality disorders in circumplex and five-factor space: explorations within a clinical sample. *Psychological Assessment* **5**, 41–52.

Stepp, S. D., Trull, T. J., Burr, R. M., Wolfenstein, M. & Vieth, A. Z. (2005). Incremental validity of the Structured Interview for the Five-Factor Model of personality (SIFFM). *European Journal of Personality* **19**, 343–357.

Stevens, J. (2002). *Applied Multivariate Statistics for the Social Sciences* (4th edn). Lawrence Erlbaum and Associates: Mahwah, NJ.

Trull, T. J. & Durrett, C. A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology* **1**, 355–380.

Trull, T. J. & Widiger, T. A. (1997). *Structured Interview for the Five-Factor Model of Personality* (*SIFFM*): *Professional Manual*. Psychological Assessment Resources: Odessa, FL.

Trull, T. J., Widiger, T. A. & Burr, R. (2001). A structured interview for the assessment of the Five Factor Model of Personality: facet-level relations to the Axis II personality disorders. *Journal of Personality* **69**, 175–198.

Tyrer, P., Seivewright, H. & Johnson, T. (2003). The core elements of neurosis: mixed anxiety-depression (cothymia) and personality disorder. *Journal of Personality Disorders* **17**, 129–138.

Warner, M. B., Morey, L. C., Finch, J. F., Gunderson, J. G., Skodol, A. E., Sanislow, C. A., Shea, M. T., McGlashan, T. H. & Grilo, C. M. (2004). The longitudinal relationship of personality traits and disorders. *Journal of Abnormal Psychology* **113**, 217–227.

Weissman, M. & Bothwell, S. (1976). The assessment of social adjustment by patient self-report. *Archives of General Psychiatry* **33**, 1111–1115.

Widiger, T. A. & Clark, L. A. (2000). Toward DSM-V and the classification of psychopathology. *Psychological Bulletin* **126**, 946–963.

Widiger, T. A., Trull, T. J., Clarkin, J. F., Sanderson, C. & Costa, P. T. (1994). A description of the DSM-III-R and DSM-IV personality disorders with the five-factor model of personality. In *Personality Disorders and the Five-Factor Model of Personality* (ed. P. T. Costa and T. A. Widiger). American Psychological Association: Washington, DC.

Zanarini, M. C., Frankenburg, F. R., Sickel, A. E. & Yong, L. (1996). *The Diagnostic Interview for DSM-IV Personality Disorders* (*DIPD-IV*). McLean Hospital: Belmont, MA.

Zanarini, M. C., Gunderson, J. G., Marino, M. F. & Schwartz, E. O. (1989). Childhood experiences of borderline patients. *Comprehensive Psychiatry* **30**, 18–25.

Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., Morey, L. C., Grilo, C. M., Shea, M. T., McGlashan, T. H. & Gunderson, J. G. (2000). The collaborative longitudinal personality disorders study: reliability of axis I and axis II diagnoses. *Journal of Personality Disorders* **14**, 291–299.