

Mixture Modelling of DSM-IV-TR Paranoid Personality Disorder Criteria in a General Population Sample

Sharon Devine¹, Brendan Bunting¹,
Siobhan McCann¹, and Sam Murphy¹

Abstract

Complications in the research into personality disorders may be rooted in the assumption within psychiatric diagnosis that underlying constructs are measured with equally valid observed items without rank or recognition of measurement error. The aim of this paper is to investigate the internal validity of DSM-IV (APA, 2000) paranoid personality disorder while accounting for measurement error and the continuous and categorical nature of the construct. General population data from the British Psychiatric Morbidity Survey (Singleton et al., 2001) was obtained from the Data Archives, University of Essex, England. Information from individuals with responses in the paranoid personality disorder section (n=8393) of the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II; First et al., 1997) screening questionnaire was analysed using confirmatory factor analysis (CFA), item response theory (IRT), latent class analysis (LCA) and latent class factor analysis (LCFA) mixture modelling. Results indicated that a one-factor model adequately represented the data, and that all items had reasonable factor loadings. However IRT analysis indicated that only four of the seven criteria discriminate well between individuals along different points of the underlying continuum. LCA and LCFA provided another perspective on the evaluation of paranoid personality disorder and indicated the presence of four underlying sub-populations. This is useful in terms of clinical and primary health settings as specific groups of interest can be investigated further in terms of characteristics, covariates and predictors.

1 Introduction

The internal validity of constructs is commonly assessed through factor analysis. However many measures in regular use have not undergone explicit testing, but are developed from experience and through the work of expert panels; such is the case

¹ Psychology Research Institute, University of Ulster, Northern Ireland; Devine-S@ulster.ac.uk

with many psychiatric diagnostic systems. Psychiatric diagnoses are based upon an analysis of manifest responses and these are then used to evaluate a latent diagnostic classification in the form of a summed index with given cut-off points. Both factor analysis and summed indices contain their own sets of assumptions.

The summed index approach assumes that one underlying construct is measured and all observed items are equally valid, relevant and without rank, weight or preference. There is no distinction made between the collection of observed items and the latent variable, and therefore measurement error (whether random or systematic) is unaccounted for. In the summed index approach it is not clear how any dimensional structure of the underlying construct is assessed. Factor analysis assumes the factor is a continuum and considers this construct in terms of continuous latent variable true values and measurement error; however this makes it difficult to identify natural cut-off points or thresholds for diagnosis.

Implicit in both factor analytic and summed index approaches is the assumption that the individuals can be usefully seen as coming from one underlying population. However, in psychiatric-type research it may be better to approach the structure in terms of sub-populations and then to establish the extent of the problem within the subpopulations of interest. Both the summed index approach and the more general model of factor analysis can be formulated in a manner that takes account of sub-populations, through the use of a mixture modelling approach to the analysis, i.e., latent class factor analysis. Methods in psychiatric research have included categorical and dimensional modelling; however as medical nomenclatures provide categorical representations of disorder and disease methods reflecting this have been prominent. Continuous representations have been used for research purposes in order to utilize constructs such as severity and to examine occurrence of sub-threshold levels of disorder or disease, leading to the increasing demand that dimensional facets are included in new editions of psychiatric nomenclatures. This is most evident in the current debate surrounding the representation of Personality Disorders, which are represented on Axis II of the current Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000). There are currently 10 specific personality disorders listed and described in this nomenclature. However, personality disorder not otherwise specified (PDNOS) is the most prevalent diagnosis (in non-structured interview studies; Verhul and Widiger, 2004), reflecting a less than optimal classification system. A number of dimensional models of personality disorder have been suggested, ranging from prototype matching approaches (for example, Oldham and Skodol, 2000; Westen and Shedler, 2000) to personality trait dimensional models (for example, Livesley and Jackson, 2004; Clark, 1999; Costa and McCrae, 1990; Cloninger et al., 1993; - see Trull and Durrett, 2005, for a complete overview of these approaches), and severity dimensions (Trull and Johnson, 1996). One of the main issues currently debated is how to integrate dimensional modelling with the current categorical system.

Traditionally in research categorical and dimensional representations assume different statistical modelling techniques. Yet some statistical techniques utilise dimensional modelling (such as factor analysis which provides a continuous representation of the underlying construct) together with categorical modelling (such as latent class analysis, in which discrete sub-populations are assumed to be present in the population). For example, within a sub-population or category there may be a continuum of severity and therefore it is appropriate to integrate the two perspectives. In a psychiatric context some subpopulations may be of particular interest and such mixture modelling is a parsimonious way of identifying the composition of sub-groups. In a recent example of this, Muthén and Asparouhov (2006) used DSM-IV (APA, 1994) tobacco dependence criteria to illustrate combination/ mixture modelling within a psychiatric diagnosis framework. The current paper investigates the internal validity of DSM (APA, 1992; 2000) paranoid personality disorder diagnostic criteria using a latent trait and mixture modelling approach that provides information on both dimensional and categorical representations of this disorder.

2 Method

Data: Data from the British Psychiatric Morbidity Survey (BPMS) of Adults Living in Private Households (2000) (Singleton et al., 2001) was obtained from the UK Data Archive, University of Essex. This dataset contains a wide range of psychiatric, psychological, physical, social and demographic information on 8580 individuals from England, Scotland and Wales. More information can be found in the survey technical report (ibid.). Information on weighting, clusters and stratification were taken into account within analyses. The paranoid personality disorder section provided details on 8393 participants, aged 16-74 years ($m=45$, $s.d.=15$ years), 44.7% were male and 55.3% were female.

DSM scoring: The Structured Clinical Interview for DSM Axis II Disorders (SCID-II; First et al., 1997) self-report screening questionnaire was used to collect information on personality disorder symptoms. Diagnostic criteria were coded as present (1) or absent (0). DSM-IV-TR diagnosis for paranoid personality disorder is indicated by the presence of any four or more of seven criteria (which can be seen in Table 1). Within the BPMS dataset, thresholds for screening positive for paranoid personality disorder were elevated to the presence of at least five of the seven criteria to match algorithms developed from an earlier study (Singleton, Meltzer, Gatward, Coid and Deasy, 1998) for concordance with clinical interview diagnosis.

Procedure: Statistical analyses were carried out using Mplus Version 4.2 software (Muthén and Muthén, 1998-2007). The modelling was developed on a step-by-step basis. 1) Given that DSM-IV-TR has brought together seven criteria that describe and measure the specific disorder of paranoid personality disorder,

confirmatory factor analysis was undertaken to test a one-factor model. Weighted Least Squares Means and Variances (WLSMV) estimation method was utilized as dichotomous items were examined. 2) To obtain further information regarding the discriminatory and severity characteristics of each criterion, item response analysis was carried out. 3) Latent class analysis was then performed to provide categorical representation of the data. 4) The final model tested was a latent class factor analysis to investigate the continuous factor at the categorical levels and to examine whether the one factor model held across the different sub-populations. The degree of concordance between these different approaches to diagnostic classification was then examined.

3 Results

The seven DSM (APA, 2000; 1992) criteria describing paranoid personality disorder and the percentage of the sample endorsing them are shown in Table 1. The range of possible criteria endorsements (summed index) was 0-7 for paranoid personality disorder.

Table 1: Percentage of sample endorsing paranoid personality disorder criteria.

DSM-IV-TR Paranoid Personality Disorder Criteria (A)	Total endorsed
1. Suspects, without sufficient basis, that others are exploiting, harming, or deceiving him or her.	28.5%
2. Is preoccupied with unjustified doubts about the loyalty or trustworthiness of friends or associates.	16.3%
3. Is reluctant to confide in others because of unwarranted fear that the information will be used maliciously against him or her.	22.8%
4. Reads hidden demeaning or threatening meanings into benign remarks or events.	19.4%
5. Persistently bears grudges, i.e., is unforgiving of insults, injuries, or slights.	29.0%
6. Perceives attacks on his or her character or reputation that are not apparent to others and is quick to react angrily or to counterattack.	18.2%
7. Has recurrent suspicions, without justification, regarding fidelity of spouse or sexual partner.	10.2%

The most popular level in this range was zero, with 3377 (40.2%) participants endorsing no criteria. The number of participants in each scoring level steadily dropped as the numbers of endorsed criteria increased: 1,944 (23.2%) endorsed one; 1,140 (13.6%) endorsed two; 787 (9.4%) endorsed 3; 545 (6.5%) endorsed 4; 349 (4.2%) endorsed 5; 184 (2.2%) endorsed 6; and 67 (0.8%) endorsed all seven criteria. Four or more criteria (DSM diagnostic threshold level) were endorsed by 13.7% (1145) of the sample; and 7.2% (600) met and/or exceeded the threshold of five endorsements (diagnostic threshold used in the survey).

Confirmatory Factor Analysis (CFA): The presence of one underlying construct of paranoid personality disorder was tested (as described above in Procedure step 1). The one-factor model adequately fitted the data, with both the Tucker Lewis Index (TLI) and the Comparative Fit Index (CFI) providing values of 0.95, RMSEA = .049. Factor loadings for the items were reasonable and ranged from 0.473 (criterion five) to 0.857 (criterion two). Factor loadings and standard errors for all criteria are presented in Table 2. The threshold for modification indices was set at 5.0, and no modification exceeded this value. Factor scores ranged from -0.44 to 1.51. When examined according to the two DSM summed index diagnostic groups (can be seen in Table 4), all individuals with factor scores ≥ 1.04 (5.1%; n=428) were in the ‘has disorder’ group, those with scores < 0.8 (87.8%; n=7370) were in the ‘no disorder’ group, leaving a factor score range of 0.8 - 1.03 (inclusive; 7.1%; n=596) in which the factor score value was not indicative of DSM diagnostic group membership.

Table 2: CFA factor loadings for paranoid personality disorder criteria (WLSMV estimation method).

Criteria descriptions (items)	Standardised factor loading	Unstandardised factor loading	SE
1. Suspicion of exploitation, harm, deception.	0.722	1.00	0.00
2. Preoccupation with doubts about loyalty and trustworthiness.	0.857	1.186	0.028
3. Fear of confiding due to malicious use of information.	0.782	1.082	0.027
4. Reads demeaning or threatening meanings.	0.777	1.076	0.029
5. Persistently bears grudges.	0.473	0.655	0.029
6. Perceives attacks on character and quick to react angrily or counterattack.	0.527	0.730	0.030
7. Recurrent suspicions regarding fidelity of spouse or sexual partner.	0.513	0.711	0.031

Item Response Theory (IRT): Having established one underlying factor, an IRT analysis using a 2-parameter logistic model was carried out to give a different perspective on the data. Although related to CFA, IRT places emphasis on the severity and discrimination characteristics of each criterion and provides graphical representations that are useful for illustration. The severity (difficulty) characteristic refers to the location or value on the underlying continuum at which there is a 50% probability of a specific response to the item (Rodebaugh et al., 2004). The discrimination characteristic refers to the how much the probability of a specific response changes for a change in value on the underlying continuum. A steep item characteristic curve (ICC) indicates good discrimination properties, as the probability increases sharply. For more information, a good introduction to IRT can be found in Baker (2001). Figure 1 shows ICCs for the seven items. The

underlying continuum of paranoid personality disorder runs along the x-axis, and severity (difficulty) characteristics are identified where the curve cuts the 0.5 probability level of endorsement (measured on y axis). Criteria one to four appear stronger for discrimination at the higher end of the factor scale. However criteria five, six and seven have very gentle slopes indicating lower discriminatory power, and severity levels for items six and seven are not reached at three standard deviations above the mean in the underlying continuum.

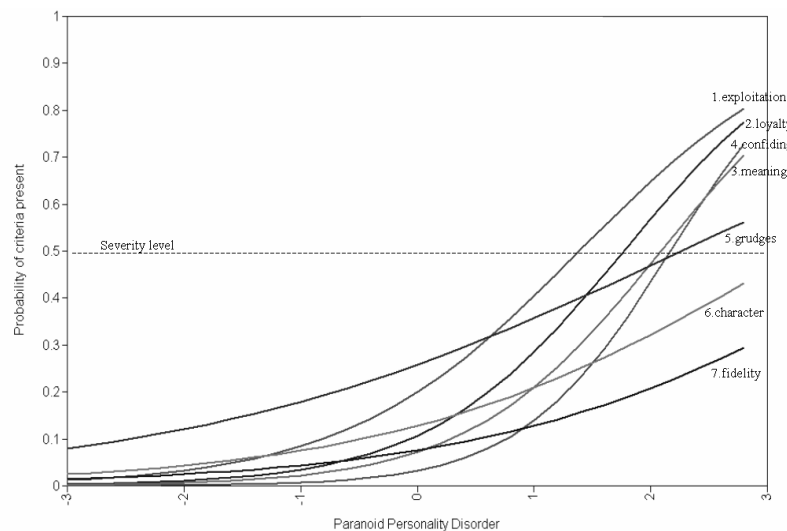


Figure 1: Item characteristic curves for the seven paranoid personality disorder criteria (MLR estimation method)².

Latent Class Analysis (LCA): Having observed the underlying discriminatory and severity characteristics of each item, the next step was to investigate the patterns of item responses within the data set and establish any sub populations. Latent class analysis examines the patterns of observed responses and divides the sample into latent homogeneous groups, providing probabilities of how individuals within each group endorse each item as well as probabilities of correct classification into each class. As this was exploratory in nature, four models were tested, ranging from a two-class to a five-class solution, with results indicating that the four-class model best fits the data as seen in Table 3. This conclusion is reached because the fit statistics BIC and SSABIC fall in value until the four-class model, and the Lo-Mendell-Rubin (Lo, Mendell and Rubin, 2001) likelihood ratio test (LRT) non-significant probability value for the five-class solution indicates that it is not a significant improvement over the model with one less class.

Latent Class Factor Analysis (LCFA): With the identification of four latent homogeneous classes, the final step was to integrate the categorical and dimensional representations into a mixture model. LCFA allows the examination of latent classes in terms of the one underlying factor established with CFA. This

² Further details on the item parameters are available from the first author.

mixture model allows for the mean of the factor to vary between classes while maintaining measurement invariance of the factor across the classes (factor loadings, variance and thresholds are constrained to be equal across the classes). Four models were tested, ranging from 2-5 class models. The four-class model provided best fit in terms of the BIC, SSABIC and LRT results. Table 3 provides fit statistics for both the latent class analyses and the mixture model analyses.

Table 3: Fit statistics for the 2-5 class models of paranoid personality disorder (MLR estimation method).

Model	Classes	LogL value	Akaike (AIC)	Bayesian BIC	SSABIC	Entropy	
LCA	2	-26018	52066	52172	52124	0.786	0.000
	3	-25856	51758	51920	51847	0.582	0.0049
	4	-25766	51595	51813	51714	0.657	0.0004
	5	-25742	51563	51837	51714	0.656	0.5844
LCFA	2	-26018	52066	52172	52124	0.786	0.000
	3	-25879	51792	51912	51858	0.596	0.000
	4	-25857	51753	51887	51827	0.655	0.0474
	5	-25854	51750	51898	51831	0.615	0.0610

From this table it can be seen that the latent class analysis four-class model provides best fit, however the inclusion of factor measurement in the mixture model allows for the consideration of further a priori information and therefore is the preferred model. Figure two is a profile plot showing the probabilities of individuals within each class endorsing the items. Class one represents 1.8% of the sample, and given the high probabilities of endorsing the criteria can be seen as the ‘disordered’ group. Class two represents 52.7% of the sample, indicating that over half of the sample has low to medium probability of endorsing the items. The majority of the individuals have been incorporated into classes two and four, the two lowest probability classes. Class 4 represents 25.6% of the sample and as there is minimal probability of endorsement of criterion five, and zero probability of endorsing the other items, this can be seen as the ‘normal’ or baseline group.

Comparisons with DSM-IV diagnosis: Both FA and IRT analyses provide scores along the underlying continuum for each individual. When compared with the screening disordered/not disordered variable as seen in Table 4 it is evident that overlap occurs. Individuals with factor scores between 0.8 and 1.03 are represented in both DSM diagnostic categories. Similarly, IRT scores between 2.01 and 2.67 are represented in both diagnostic categories. Posterior probabilities were used to classify individuals according to their most likely class membership. The latent classes were also examined in terms of DSM diagnosis. The individuals in the two classes with lowest probabilities of endorsing the items all have a diagnosis of ‘no disorder’, all those in the class with highest probabilities have a diagnosis of ‘has disorder’, and those in the class with the second highest probabilities are represented in both diagnostic groups. Table 4 provides a cross

tabulation of the latent classes (individuals are classified into classes based on most likely class membership) with DSM diagnosis of paranoid personality disorder.

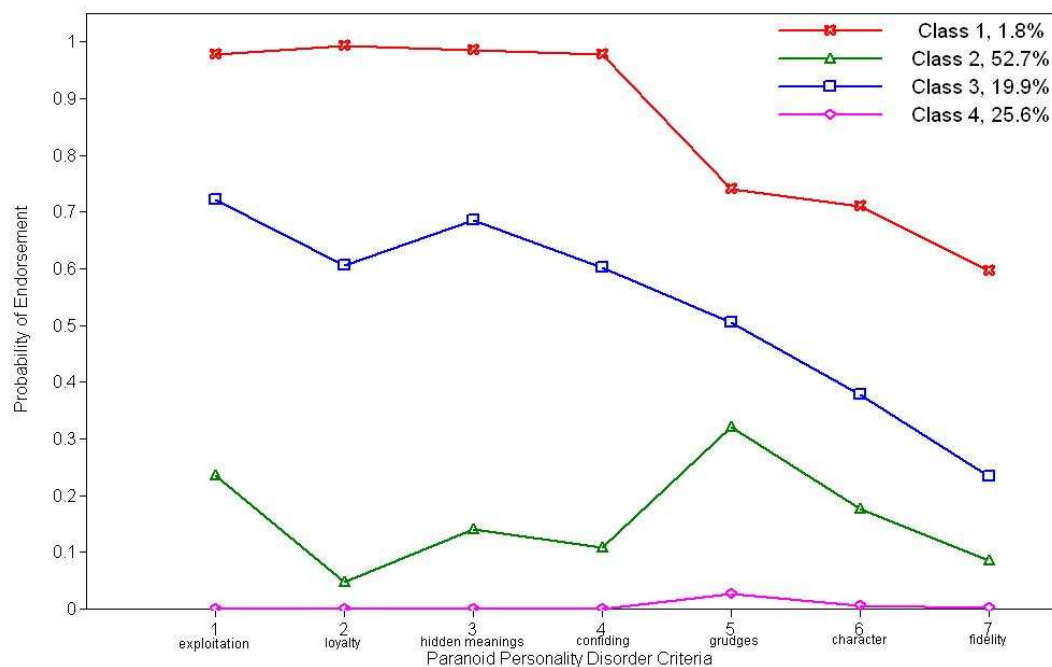


Figure 2: Latent Class Factor Analysis class profiles.

Table 4: DSM paranoid personality disorder diagnosis cross tabulation with four LCFA latent classes, factor scores and IRT scores.

N=8393		Most likely class membership				Range; M; SD	
		High Endorsement. (Class 1)	Med-High Endorsement. (Class 3)	Low-Med Endorsement. (Class 2)	No Endorsement. (Class 4)	Factor Score	IRT Score
DSM diagnosis	No disorder	0% (n = 0)	13.4% (n = 1126)	39.5% (n = 3319)	39.9% (n = 3348)	-0.44 – 1.03; -0.02; 0.44	-1.34 – 2.67; -0.23; 1.19
	Has disorder	0.9% (n = 78)	6.2% (n = 522)	0% (n = 0)	0% (n = 0)	0.80 – 1.51; 1.17; 0.19	2.01 – 3.99; 3.02; 0.51
Factor score	Range	1.34 – 1.51	0.53 – 1.32	-0.14 – 0.57			
	Mean, SD	1.48, 0.07	0.86, 0.23	0.14, 0.21	-----		
IRT score	Range	3.53 – 3.99	1.30 – 3.45	-0.59 – 1.36			
	Mean, SD	3.90, 0.19	2.21, 0.62	0.17, 0.59	-----		

4 Summary and conclusions

Dimensional and categorical representations of paranoid personality disorder data were examined and compared with the DSM summed index approach. Factor analysis indicated that a one-factor model fits the data, with factor loadings which reflect reasonable relationships in a factor analysis framework. The analysis identified a range of factor scores which may be useful in terms of research into aspects of threshold and sub-threshold presentations of paranoid personality disorder. IRT analysis provided information on two parameters – discrimination and severity – and illustrated that items one to four discriminate best, however the final three criteria do not differentiate well between individuals, and overall discrimination and severity properties of these three items indicate they are not optimal markers across the paranoid personality disorder continuum. This has ramifications for the summed index approach as it is assumed that all items are of equal value in the overall construct, and that numbers of criteria endorsed are indicative of disorder rather than identifying which items are likely to be endorsed by individuals with high levels of the disorder. In clinical terms it would be useful to have items that differentiate individuals along the continuum. Examination of the responses from individuals with factor scores in the overlapping range mentioned above may reveal that the final three criteria contribute to the lack of diagnostic discrimination.

Latent class analysis identified four sub-populations, indicating considerable heterogeneity in the population. The factor model was reanalysed in terms of these sub-populations and four latent classes were established as a good description of the data. Three of the four classes clearly differentiated between the DSM diagnostic categories of ‘disorder’ and ‘no disorder’. In the remaining class there was considerable overlap in DSM diagnosis. Further analysis is required to gain a better understanding of the degree of misclassification within this sub-population. Latent class profile plots show that criteria five, six and seven have the least probabilities of endorsement by the two highest endorsing groups, and criteria five is the most likely item to be endorsed by the two lowest endorsing groups, supporting the IRT indication that discrimination is poor in these items.

Limitations of this study include the self-report method of data collection and the extreme skewness of item endorsement. The personality disorder information was derived using a self-report measure, which may have been influenced by recall bias, both of the validity and accuracy of the responses. Clinical and informant interviews would be desirable but are unsuitable for large epidemiological surveys. Epidemiological data is useful for examination of ‘normal’ levels of clinical symptoms, however this may lead to problems with the distribution of responses. The distribution of item endorsement is heavily skewed, with almost half of the sample endorsing zero items. This may have ramifications for some of the techniques used. Individuals endorsing zero items and endorsing all items provide

large amounts of invariance which may also affect analyses, although these individuals appeared to be accounted for within LCA and LCFA profiles.

Overall, although the factor analysis of the items provided reasonable factor loadings, IRT analysis indicated poor discrimination in items used in clinical decision-making guidelines. The use of latent class analysis, while maintaining factor structure, provided the opportunity to examine the construct in a joint continuous and categorical fashion and identified four underlying populations within the factor continuum. This has clinical and primary health care utility because any sub-population of interest can be examined further in terms of characteristics which may identify and verify risk or resilience factors, or the effects of clinical interventions.

Acknowledgement

The authors would like to thank the referees for their valuable comments and suggestions.

References

- [1] American Psychiatric Association (1994): *Diagnostic and Statistical Manual of Mental Disorders, 4th ed.* Washington: Author.
- [2] American Psychiatric Association (2000): *Diagnostic and Statistical Manual of Mental Disorders, Revised 4th ed.* Washington: Author.
- [3] Baker, F.B. (2001): *The Basics of Item Response Theory.* 2nd ed. USA: ERIC Clearinghouse on Assessment and Evaluation.
- [4] Clark, L.A. (1999): Dimensional approaches to personality disorder assessment and diagnosis. In Cloninger, C.R. (Ed.): *Personality and Psychopathology* (pp. 219-244). Washington: American Psychiatric Press.
- [5] Cloninger, C.R., Svrakic, D.M., and Przybeck, T.R. (1993): A psychobiological model of temperament and character. *Archives of General Psychiatry*, **50**, 975-990.
- [6] Costa, P.T. and McCrae, R.R. (1990): Personality disorders and the five-factor model of personality. *Journal of Personality Disorders*, **4**, 362-371.
- [7] First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., and Benjamin, L. (1997): *Structured Clinical Interview for DSM-IV Axis II Personality Disorders.* Washington: American Psychiatric Press.
- [8] Livesley, W.J. and Jackson, D.N. (2004): *Dimensional Assessment of Personality Pathology-Basic Questionnaire.* Port Huron, MI: Research Psychologists Press.

-
- [9] Lo, Y., Mendell, N.R., and Rubin, D.B. (2001): Testing the number of components in a normal mixture. *Biometrika*, **88**, 767-778.
- [10] Muthén, B. and Asparouhov, T. (2006): Item response mixture modelling: Application to tobacco dependence criteria. *Addictive Behaviors*, **31**, 1050-1066.
- [11] Muthén, L.K. and Muthén, B.O. (1998-2007): *Mplus User's Guide*. Fourth Edition. Los Angeles: Muthén & Muthén.
- [12] Oldham, J.M. and Skodol, A.E. (2000): Charting the future of Axis II. *Journal of Personality Disorders*, **14**, 17-29.
- [13] Rodebaugh, T.L., Woods, C.M., Thissen, D.M., Heimberg, R.G., Chambless, D.L., and Rapee, R.M. (2004): More information from fewer questions: The factor structure and item properties of the original and brief Fear of Negative Evaluation Scale. *Psychological Assessment*, **16**, 169-181.
- [14] Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., and Meltzer, H. (2001): *Psychiatric Morbidity Among Adults Living in Private Households, 2000*. London: Her Majesty's Stationery Office (HMSO).
- [15] Singleton, N., Meltzer, H., Gatward, R., Coid, J., and Deasy, D. (1998): *Psychiatric Morbidity Among Prisoners in England and Wales*. London: The Stationery Office.
- [16] Trull, T.J. and Durrett, C.A. (2005): Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology*, **1**, 355-380.
- [17] Tryer, P. and Johnson, T. (1996): Establishing the severity of personality disorder. *American Journal of Psychiatry*, **153**, 1593-1587.
- [18] Verhul, R. and Widiger, T.A. (2004): A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders*, **18**, 309-319.
- [19] Westen, D. and Shedler, J. (2000): A prototype matching approach to diagnosing personality disorders: toward DSM-V. *Journal of Personality Disorders*, **14**, 109-126.