Psychological Medicine, 1998, **28**, 145–152. Printed in the United Kingdom

The comparison of latent variable models of non-psychotic psychiatric morbidity in four culturally diverse populations

K. S. JACOB,[1] B. S. EVERITT, V. PATEL, S. WEICH, R. ARAYA AND G. H. LEWIS

From the Section of Epidemiology and General Practice and Department of Biostatistics and Computing, Institute of Psychiatry, London and Division of Psychological Medicine, University of Wales College of Medicine, Cardiff; and Epidemiology and Community Psychiatry Unit, Department of Psychiatry, Universidad de Chile, Santiago, Chile

ABSTRACT

Background. Factor analysis has been employed to identify latent variables that are unifying constructs and that parsimoniously describe correlations among a related group of variables. Confirmatory factor analysis is used to test hypothesized factor structures for a set of variables; it can also, as in this paper be used to model data from two or more groups simultaneously to determine whether they have the same factor structure.

Method. Non-psychotic psychiatric morbidity, elicited by the Revised Clinical Interview Schedule (CIS-R), from four culturally diverse populations was compared. Confirmatory factor analysis was employed to compare the factor structures of CIS-R data sets from Santiago, Harare, Rotherhithe and Ealing. These structures were compared with hypothetical one and two factor (depression–anxiety) models.

Results. The models fitted well with the different data sets. The depression–anxiety model was marginally superior to the one factor model as judged by various statistical measures of fit. The two factors in depression–anxiety model were, however, highly correlated.

Conclusions. The findings suggest that symptoms of emotional distress seem to have the same factor structure across cultures.

INTRODUCTION

Classification of non-psychotic psychiatric morbidity has been a source of perennial controversy (Tyrer, 1985). Categorical and dimensional models have been examined (Goldberg & Huxley, 1992) and various statistical techniques have been employed to support these approaches. Categorical models have been preferred on practical considerations. However, dimensional representations are said be more accurate reflections of clinical reality (Feinstein, 1985). A number of dimensional models have been proposed to explain the relationship between depression and anxiety. Most such models involve the concept of 'latent variables' (often also referred to as components, common factors or factors), these being unifying constructs that characterize responses to related groups of variables. Latent variables, though not directly quantifiable, are derived from the measurement of other, directly observable variables. The identification of such underlying dimensions or factors greatly simplifies the description and understanding of complex phenomena, such as non-psychotic psychiatric morbidity.

Studies have most often employed principal component analysis to identify factors that parsimoniously describe the correlations between observed depression and anxiety variables.

* Address for correspondence: Dr K. S. Jacob, Section of Epidemiology and General Practice, Institute of Psychiatry, Denmark Hill, London SE5 8AF.

Research from primary care found that two dimensions provided a slightly better fit than a one dimensional solution (Goldberg *et al*. 1987). The two dimensions were not independent and were found to correlate with one another. Studies of psychiatric patients have also suggested similar two factor solutions (Carney *et al*. 1965; Hamilton *et al*. 1967; Rosenthal & Gudeman, 1967; Paykel *et al*. 1971; Kiloh *et al*. 1972). A review of studies that employed principal component analysis of symptoms, personality or illness features in patients with affective disorders found that most solutions consisted of factors indicative of anxiety and depression (Mullaney, 1984).

Exploratory factor analysis studies, however, have been criticized on many grounds (Dunn *et al*. 1993*a*). A general problem is that there is too much subjectivity in the comparison of the observed factor structure with the hypothetical structure. Most of the studies referred to above obtained the factor structure from the data and then made subjective statements about how closely the results matched the hypothesis. Other limitations include the fact that the technique is not hypothesis driven, that small samples result in unstable latent variables and that the technique is very sensitive to the psychometric properties of the scales employed.

Confirmatory factor analysis (CFA) (Cole, 1987; Bentler & Stein, 1992; Dunn *et al*. 1993*b*) on the other hand, allows for the testing of various hypotheses about the likely structure of latent variables in the data. It may also be employed to examine the possibility that different data sets have a similar factor structure. Unlike exploratory factor analysis, which does not attempt to test a specified hypothesis on the factor structure of the observed data, CFA is employed to assess a particular hypothesized factor model. When used to examine the fit of the models across different data sets, it largely eliminated the subjectivity that is inherent in 'eyeballing' of data.

Cross-cultural comparisons of mental disorders have also been criticized on other grounds (Kleinman, 1980, 1987; Littlewood, 1990). The assumption that Western diagnostic categories are culture free entities has been challenged. The 'category fallacy' argument (which criticizes the use of a nosological category developed for a particular cultural group that is then applied to members of another culture for whom it lacks coherence and whose validity is not established) has been employed to discredit comparisons. The differentiation of *etic* and *emic* perspectives of illness is emphasized. *Etic* models employ external perspectives whereas *emic* models elicit indigenous points of view on the conceptualization of the sickness episode.

The effect of observer bias in clinical assessment is also relevant. Most standardized psychiatric interviews place much emphasis on the use of clinical judgement in assessing psychiatric disorders. It has been argued that encouraging the use of clinical judgements may introduce or exaggerate observer bias (Lewis, 1991).

This paper compares data on non-psychotic psychiatric morbidity data from four different primary-care populations. The aim was to test the hypothesis that the different data sets had the same factor structure and to explore possible cross-cultural differences in neurotic symptomatology. Confirmatory factor modelling was carried out in order to assess whether all these groups had the same factor structure. The study employed data from a standardized self-report interview to reduce observer bias. It used symptom data rather than diagnostic grouping to reduce the possibility of 'category fallacy'.

METHOD

Data sets

Data from four different investigations that examined psychiatric morbidity in primary care were employed. These studies were carried out in Santiago (Chile), Harare (Zimbabwe), and in the United Kingdom. The Chilean sample was recruited in Santiago from a primary health care clinic serving the large working class urban area of Lo Prado. Consecutive attenders at the clinic, less than 65 years of age, were screened with two self-administered questionnaires and interviewed to ascertain possible psychiatric morbidity with the Spanish version of the Revised Clinical Interview Schedule (CIS-R) (Lewis *et al*. 1992; R. Araya, unpublished data).

Subjects were recruited in Harare (Zimbabwe) from two primary health care facilities and 10 traditional healer clinics in high density suburbs of the city (Patel *et al*. 1997). Consecutive attenders aged between 16 and 65 years, without

acute medical illness were included in the study. They were interviewed using the Shona version of the CIS-R.

In the United Kingdom, White subjects and ethnic Indians (from the subcontinent) living in London were studied. The White sample was recruited from Rotherhithe (South London) (Weich *et al*. 1996). Consecutive attenders aged 16 to 65 years were recruited at randomly selected surgeries in a health centre. All subjects completed a brief screening questionnaire and were also given the computer-administered CIS-R.

The Indian subjects lived in Ealing (Jacob *et al*. 1996). Consecutive women attenders at the general practice clinic were screened using the 12-item General Health Questionnaire. All individuals who scored $\geqslant 2$ and an equal number of those who scored 0 or 1 were interviewed using the CIS-R.

The number of subjects included in the analysis from Santiago, Harare, Rotherhithe and Ealing were 1700, 302, 273 and 100 respectively. All subjects from the United Kingdom were general practice attenders. The Santiago and the Harare samples employed individuals attending primary-care facilities. All four investigations had similar aims of identifying non-psychotic psychiatric morbidity in primary care.

Ascertainment of psychiatric morbidity

All studies employed the same method for identification of psychiatry morbidity using the Revised Clinical Interview Schedule (CIS-R) (Lewis *et al*. 1992). This instrument is a standard semi-structured interview to assess the mental state of individuals with non-psychotic psychiatric disorders. The instrument does not employ the interviewers clinical judgement and hence minimizes observer bias. Therefore, many aspects of interviewing style are prescribed by the interview including the exact wording of most of the questions and specific rules for coding each symptom. Definitions are standardized and clinical judgement replaced by rules. This also makes it ideal for use by 'lay' interviewers. The CIS-R limits detailed enquiry to the previous week on grounds that memory for psychological symptoms, and thus for the validity of responses, becomes poor when a longer period of enquiry is used.

The instrument has 14 subsections, which include somatic symptoms, fatigue, concentration, sleep problems, irritability, worry about physical health, depression, depressive ideas, worry, anxiety, phobia, panic, obsessions and compulsions. Scores for subsections range from 0–4 (0–5 for depressive ideas). The subsection employed to elicit some of the depressive symptoms is listed in Appendix 1.

The ratings obtained at interview can be presented for each symptom group and they can be summed to yield an overall score that is taken to indicate the severity of any minor psychiatric disorder. The subject is also asked about the duration of symptoms. Algorithms have also been developed, which allow an ICD-10 diagnosis. For example information from the CIS-R can be employed to make a range of ICD-10 diagnoses including mild, moderate and severe depressive episodes, agoraphobia, social phobia, panic disorder, generalized anxiety disorder, obsessive–compulsive disorder and neurasthenia. The CIS-R has been documented to have high inter-rater reliability (Lewis *et al*. 1992). It has been employed in many investigations attempting to study non-psychotic psychiatric morbidity.

The CIS-R was translated into Spanish, Shona and Hindi for use in Santiago, Harare and in Ealing respectively. The process of translation and back translation was done by bilingual mental health professionals. The process of translation of the CIS-R into different languages did not pose difficulties as the instrument is designed for use by lay interviewers and hence does not contain abstract psychiatric concepts. Rigorous translation with emphasis on conceptual equivalence and use of standard translation–back translation and consensus agreement ensured that the translated versions were valid for use in that language. In addition, the stem questions employed for screening contained many phrases describing the various symptoms. Such as design made its translation and use into other languages easier and resulted in related symptoms being identified with a single stem question. Interviewers in all studies were trained in the use of the instrument. The training included video demonstrations and supervised interviews. Inter-rater reliability for the vernacular versions was also high (Lewis *et al*. 1992; Patel *et al*. 1995).

Statistical analysis

Exploratory factor analysis is a statistical technique that attempts to identify a relatively small number of factors (or latent variables) that account for the relationships among sets of observed variables. Confirmatory factor analysis (CFA), on the other hand, tests whether a particular factor structure is consistent with the correlations or covariances of a set of observed variables (Cole, 1987; Bentler & Stein, 1992; Dunn *et al.* 1993*b*). CFA is often used in social and behavioural sciences. The aim of fitting such models is two fold: (*i*) to test some particular theory; and (*ii*) to produce a parsimonious description of data. CFA can be used in modelling two or more groups simultaneously to determine to what extent the group covariance or correlation matrices have the same structure in terms of a particular factor analysis model. If exactly the same model is found not to be appropriate for all groups, then attempts can be made to allow some terms in the model (e.g. factor correlations) to differ to achieve a more reasonable fit.

CFA differs from exploratory factor analysis (including principal component analysis) in that the investigator must specify a factor model before the data is analysed. In essence, this means the structure of the latent variables in terms of the observed variables has to be specified explicitly. The loadings on some observed variables will be fixed at zero with other loadings being free to be estimated. In addition, correlations between latent variables might be fixed at zero or allowed to be free. The free parameters in a CFA are estimated by making the predicted correlations of the variables as 'close' as possible to their observed values (see Dunn *et al.* 1993*b* for details). The 'fit' of a model can be examined in a variety of ways as described by Dunn *et al.* (1993*b*). Here we shall use the various 'fit indices' that have been developed including the normed fit index (NFI), the non-normed fit index (NNFI) and the comparative fit index (CFI). A value of 0·9 or more is considered an adequate fit. The NNFI has the advantage of reflecting model fit well at all sample sizes. In addition to these fit indices the chi-squared statistic and the residual covariance or correlation matrix also reflect the goodness-of-fit. Smaller chi-squared values and residuals suggest better fit. However, *P* values based on the chi-squared test are not usually very helpful as an indicator of fit particularly when large samples are involved. If models provide a good fit they are retained and considered plausible. Other theoretically feasible alternatives are also tested.

The EQS modelling package (Bentler, 1993) was used to test various factor models using correlation matrices of the four samples. A multi-group analysis, which analysed the four data sets as separate samples, was carried out. The EQS was used to assess similarities and differences of factor structures in the four groups. A one factor model (in which all sections of the CIS-R were loaded as a single factor) and the two factor depression–anxiety model were studied. These models were fitted to all four study groups. The two factor model tested the notion that there were two latent variables of depression and anxiety. The CIS-R subsections for worry, anxiety, phobia, panic, obsessions, compulsions used to specify the 'anxiety dimension' while the other sections (somatic symptoms, fatigue, concentration, sleep problems, irritability, worry about physical health, depression, depressive ideas) constituted the 'depression dimension' of the depression–anxiety model. However, the low prevalence of phobia, panic, obsessions and compulsions in all data sets constrained the correlations and were consequently omitted.

The first step was to test one and two factor models assuming that all four samples had identical parameters (i.e. factor loading, error terms and between factor correlations (for the two factor model)). Subsequently, the Lagrange multiplier (LM) test was employed to indicate which of these constraints it would be most advantageous to release (Dunn *et al.* 1993*b*). The LM test, evaluates the statistical necessity of the restrictions imposed by the model. The test is useful for evaluating whether, from a statistical point of view, the model could be improved substantially by freeing a previously fixed parameter. For each fixed parameter the test will give an estimate of the value that the parameter might take if it were to be freely estimated rather than constrained. The test also gives the change in the chi-squared test statistic that would result from freeing each fixed parameter.

Ideally, some general requirements have to be

met before CFA can be used with confidence to investigate a particular model (Cole, 1987). First is the question of sample size. Measures of fit and standard errors in CFA are based on large sample theory, which means that results may be misleading when sample size is small. Some relatively *ad hoc* suggestions have been made about sample size, for example, for relatively simple data sets (i.e. one, two or three factors) a minimum of 100 subjects is required (Bearden *et al.* 1982). Such 'rules' are probably sensible to keep in mind but are not cast in tablets of stone. Secondly, in a single factor model, at least three measures of the factor are required. In many multiple factor models, two measures per factor may be sufficient. The theory behind CFA also properly requires the data to have a multivariate normal distribution. Departures from this assumption may affect in particular the chi-squared measure of fit; in general, however, most applications of the technique have largely ignored this requirement, appealing perhaps to a general confidence in the robustness of most normally based statistical procedures.

RESULTS

The description of the different samples can be obtained from the original references (Jacob *et al.* 1996; Weich *et al.* 1996; Patel *et al.* 1997; R. Araya, unpublished data). The proportion of individuals who received symptom scores of > 2 in each sample are recorded in Table 1. Symptoms of phobia, panic, obsessions and compulsions had a low prevalence across all data sets are were omitted from the analysis.

CFA was employed to fit a one factor model. The fit indices of the one factor model were reasonable (Table 2). When all constraints were equal across the four samples, the non-normed and the comparative fit indices were greater than 0·9 suggesting a good fit. The Lagrange multiplier test was employed to test constraints for the loadings and error terms for all variables. Significant values for the LM test were obtained for the loadings for variables worry, anxiety and concentration (implying that the model could be improved substantially by freeing these previously fixed parameters). Consequently, the constraints on these loadings were removed and they were allowed to differ. The models were then refitted separately for all four study groups using the EQS. The goodness-of-fit improved when the constraints on the leadings for the observed variables worry and anxiety were released. When the constraints for the loadings for the variables worry, anxiety and concentration were released simultaneously a further improvement was recorded for the normed fit index.

Table 1. *Number of individuals with symptom scores of* > 2

	Santiago *N* (%)	Rotherhithe *N* (%)	Ealing *N* (%)	Harare *N* (%)
Somatic symptoms	711 (41·8)	58 (21·2)	25 (25·0)	99 (32·7)
Fatigue	728 (42·8)	109 (39·9)	38 (38·0)	87 (28·8)
Concentration	291 (17·1)	50 (18·3)	11 (11·0)	26 (8·6)
Sleep	538 (31·6)	51 (18·7)	20 (20·0)	55 (18·3)
Irritability	452 (26·5)	50 (18·3)	12 (12·0)	36 (11·9)
Preoccupation with physical health	417 (24·5)	26 (9·5)	18 (18·0)	94 (31·2)
Depressive symptoms	489 (28·7)	43 (15·8)	21 (21·0)	82 (27·1)
Depressive ideas	201 (11·8)	78 (28·6)	15 (15·0)	76 (25·2)
Worry	630 (37·1)	71 (26·0)	19 (19·0)	107 (35·5)
Anxiety	748 (44·0)	42 (15·4)	1 (1·0)	110 (36·4)
Phobia	72 (4·2)	8 (3·0)	2 (2·0)	4 (1·3)
Panic	145 (8·5)	21 (7·7)	0 (0·0)	28 (9·3)
Obsessions	154 (9·5)	54 (19·8)	0 (0·0)	2 (0·6)
Compulsions	99 (5·8)	19 (7·0)	0 (0·0)	1 (0·3)

Table 3 documents the indices for the two factor depression–anxiety models fitted on all four sets of CIS-R data. The fit indices were good for this model when all loadings and error terms were constrained to be equal across data sets. Significant values for the LM test were again obtained for loadings for the variables worry, anxiety and concentration. Consequently, the constraints on these were removed and they were allowed to differ. The models were then refitted. The goodness-of-fit improved further with the release of constraints for loadings for the variables worry, anxiety and concentration. The fit indices for the two factor models was marginally better than that obtained by the corresponding one factor models. The factors in the two factor models had an estimated correlation of 0·89.

Other models including a three factor model with somatic, depression and anxiety factors were fitted. However, the fit was poor with fit indices of less than 0·9.

https://doi.org/10.1017/S0033291797005710 Published online by Cambridge University Press

Table 2. *Indices for one factor models*

Constraints	χ^2 value	df	Normed fit index	Non-normed fit index	Comparative fit index
Multi-group analysis with all constraints equal	911*	200	0·87	0·91	0·90
Multi-group analysis with all constraints equal except worry, anxiety	779*	188	0·89	0·92	0·92
Multi-group analysis with all constraints equal except worry, anxiety, concentration	757*	182	0·90	0·92	0·92

* $P < 0{\cdot}001$.

Table 3. *Indices for two factor models*

Constraints	χ^2 value	df	Normed fit index	Non-normed fit index	Comparative fit index	Correlation
Multi-group analysis with all constraints equal	860*	199	0·88	0·93	0·91	0·89
Multi-group analysis with all constraints equal except worry, anxiety	722*	187	0·90	0·92	0·92	0·89
Multi-group analysis with all constraints equal except worry, anxiety, concentration	701*	181	0·90	0·93	0·93	0·89

* $P < 0{\cdot}001$.

DISCUSSION

This is the first study, of which we are aware, to use confirmatory factor analysis in the assessment of models for non-psychotic psychiatric morbidity. We have used it here to examine the possibility that all four samples had a similar factor structure and to test specific factor models. The number of subjects in the different samples is adequate, the data sets were complete and multiple measures were employed for each factor. The criterion for multivariate normalcy was not satisfied. However, specific factor loadings are less affected by this violation.

Methodological issues related to transcultural research include the 'category fallacy' argument which is often employed to invalidate cross-cultural comparisons. However, the CIS-R data employed in this analysis is information on symptoms of non-psychotic psychiatric morbidity and did not employ any form of diagnostic categorization. The use of symptoms rather than diagnostic categories for comparison lessens the possibility that such a fallacy was the reason for similarity in factor structure. The CIS-R, a self-report interview, reduces the possibility of observer bias by Western trained interviewers. In addition, the multiple phrases (to describe symptoms) used in the stem questions for screening allowed for related behaviours to be identified. However, the CIS-R was developed in the West and employs European perspectives on the type of symptoms associated with emotional distress. It does not take into account symptoms of non-psychotic psychiatric morbidity specific to other cultures. Nevertheless, the symptoms employed in this analysis were those which were commonly prevalent in all four populations and symptoms with low prevalence were excluded. In addition, it has been demonstrated that the symptoms questionnaires employed to recognize common mental disorders that were specifically developed within particular cultures (e.g. Shona Symptom Questionnaire) (Patel *et al.* 1997) have marked similarities in content to instruments developed in the West. Similarly, marked overlap of *etic* and *emic* concepts of common mental disorder have been described (Patel *et al.* 1995; Jacob *et al.* 1996). Consequently, this study attempted to test the universality of the factor structure of the symptoms frequently experienced across cultures. The findings argue for the universality of the experience of symptoms of emotional distress and a similar relationship between symptoms.

The one factor model hypothesized that the scores on the observed symptom variables were

all reflecting the subject's morbidity on a single dimension. The CFA tested whether all symptoms could be parsimoniously reduced to a single latent variable. The analysis demonstrated that the one factor model fitted well in the Santiago, Harare, Rotherhithe and Ealing samples.

Despite the good fit of the one factor model attempts were made to identify cross-cultural differences. The variances and regression coefficients of variables that differed significantly across data sets were identified using the LM test and included concentration, worry and anxiety. The one factor model was then refitted allowing these to differ across populations while holding the others constant. The refitting of this model after the release of their constraints improved the goodness-of-fit. This suggests that while the most observed variables seemed to be similar across samples, the observed variables for worry, anxiety and concentration differed across populations and releasing these constraints resulted in an improved fit of the model.

The two factor depression–anxiety model was also similarly tested. The model also fitted the observed data well. Permitting the variances and coefficients of variables worry, anxiety to vary across data sets improved the fit suggesting that the 'depression' factor is similar across cultures while there are differences in the 'anxiety' factor. The fit indices of the depression–anxiety model were marginally superior to the corresponding one factor models. However, the high estimated correlation of about 0·9 between the two factors strongly suggests that the two latent variables are not independent and that consequently, high levels of depression are associated with high levels of anxiety and vice versa. The fact that the one factor model (where all symptoms loaded on to a single factor) fitted equally well in comparison to a two factor model argues that the depression–anxiety division may be artificial and that all symptoms can be explained by a single unifying construct. Other investigations have also reported similar findings (i.e. marginal superiority of the two factor solution over the one factor model and the high correlation between the two factors) (Goldberg *et al*. 1987). One could also argue that the depression–anxiety distinction is derived from Western tradition and less appropriate for cross-cultural work. The one factor model does not have any of the Western theoretical assumptions regarding the grouping of different symptoms. It argues that all symptoms reflect the subject's morbidity and that they can be parsimoniously reduced to a single latent variable. For these reasons the one factor model may be preferred to the two factor solution in primary-care settings. The one factor model also supports the approach of most research in community settings that treat psychiatric morbidity as a single dimension. More work needs to be done in this domain to confirm these findings.

A criticism of this study would include the apparent circular logic of choosing symptoms commonly seen in all cultures and then demonstrating similarity in factor structures. However, factor structures are dependent more on the correlations between symptoms than on their actual prevalence. The similarity in factor structure of the symptoms of non-psychotic psychiatric morbidity, despite the varied concepts of mental illness seen across cultures, suggests universality of the relationship among symptoms.

Conclusion

Confirmatory factor analysis was employed to compare the factor structures of CIS-R data sets from Santiago, Harare, Rotherhithe and Ealing. These structures were compared with hypothetical one and two factor models. The models fitted well with the different data sets. The fit improved when the variances and coefficients of some observed variables were allowed to differ across data sets. The findings suggest that the experience of symptoms of emotional distress seems to have the same factor structure across cultures.

K. S. Jacob is a Wellcome Fellow. V. Patel was supported by the Beit Medical Fellowship and the International Development Research Centre (Canada).

APPENDIX 1

The subsection of the CIS-R employed to elicit 'depressive' symptoms

Mandatory screening questions

1 'Almost everyone becomes low in mood or depressed at times. Have you had a spell of feeling sad, miserable or depressed in the past month?'

2 'During the past month, have you been able to enjoy or take an interest in things as much as you usually do?'

Confirmatory questions

(These were asked to individuals who answered positively to either screening question. The interviewer is instructed to use the informants own words, if possible, to ask these questions.)
1 'In the past week, have you had a spell of feeling sad, miserable or depressed?'
2 'In the past week, have you been able to enjoy or take an interest in things as much as usual?'
3 'In the past week, on how many days have you felt sad, miserable or depressed or unable to enjoy or take an interest in things?'
4 'Have you felt sad, miserable or depressed or unable to enjoy or take an interest in things for more than 3 hours in total on any day in the past week?'
5 'What is the MAIN thing that made you feel sad, miserable or depressed or unable to enjoy or take an interest in things in the past week? Can you choose from this card?' (The list included: family members, including spouse or partner; relationships with friends or people at work; housing; money or bills; your own physical health, including pregnancy; you own mental health; work or lack of work (including studying); legal difficulties; political issues or the news; other.)
6 'In the past week when you felt sad, miserable or depressed or unable to enjoy or take an interest in things, did you ever become happier when something nice happened, or when you were in company?'
7 'How long have you been feeling sad, miserable or depressed or unable to enjoy or take an interest in things as you have described?'
8 Subsections of the CIS-R were also employed to assess the depression dimension and included sections for somatic complaints, fatigue, problems in concentration, worry over physical health, irritability and sleep problems. The subsection elaborated in the Appendix assessed depressed mood/affect. The section on depressive ideas assessed symptoms like sleep, appetite, change in weight, interest in sex, diurnal variation of mood, guilt, restlessness, motor retardation, hopelessness and suicidal ideation.

REFERENCES

Bearden, W. O., Sharma, S. & Teel, J. E. (1982). Sample size effects on chi-square and statistics used in evaluating causal models. *Journal of Marketing Research* **19**, 425–430.

Bentler, P. M. (1993). *EQS/Window's User Guide*, BMDP Statistical Software, Inc.: Los Angeles.

Bentler, P. M. & Stein, J. A. (1992). Structural equation models in medical research. *Statistical Methods in Medical Research* **1**, 159–181.

Carney, M. N. P., Roth, M. & Garside, R. F. (1965). The diagnosis of depressive syndromes and prediction of ECT response. *British Journal of Psychiatry* **111**, 659–674.

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology* **55**, 584–594.

Duan, G., Sham, P. & Hand, D. (1993*a*). Statistics and the nature of depression. *Psychological Medicine* **23**, 871–889.

Dunn, G., Everitt, B. & Pickles, A. (1993*b*). *Modelling Covariances and Latent Variables using EQS*. Chapman & Hall: London.

Feinstein, A. R. (1985) *Clinical Epidemiology: The Architecture of Clinical Research*. WB Saunders Company: Philadelphia.

Goldberg, D. & Huxley, P. (1992). *Common Mental Disorders: A Bio-social Model*. Routledge: London.

Goldberg, D. P., Bridges, K., Duncan-Jones, P. & Grayson, D. (1987). Dimensions of neuroses seen in primary-care settings. *Psychological Medicine* **17**, 461–470.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* **6**, 278–296.

Jacob, K. S., Bhugra, D., Lloyd, K. R. & Mann, A. H. (1996). Psychiatric morbidity, explanatory models and consultation behaviour among ethnic Indians living in the United Kingdom. In, Abstracts of the Eighth Symposium of the Association of European Psychiatrists: Section of Epidemiology and Social Psychiatry, 11–13 April, 1996, Cambridge.

Kiloh, L. G., Andrews, G., Neilson, M. & Bianchi, G. N. (1972). The relationship of syndromes called endogenous and neurotic depression. *British Journal of Psychiatry* **121**, 183–196.

Kleinman, A. (1980). *Patients and Healers in the Context of Culture: An Exploration of the Borderland between Anthropology, Medicine, Psychiatry*. University of California Press: Berkeley, CA.

Kleinman, A. (1987). Anthropology and psychiatry: the role of culture in cross-cultural research on illness. *British Journal of Psychiatry* **151**, 447–454.

Lewis, G. (1991). Observer bias in the assessment of anxiety and depression. *Social Psychiatry and Psychiatric Epidemiology* **26**, 265–272.

Lewis, G., Pelosi, A., Araya, R. & Dunn, G. (1992). Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychological Medicine* **22**, 465–486.

Littlewood, R. (1990). From categories to contexts: a decade of the 'new cross-cultural psychiatry'. *British Journal of Psychiatry* **156**, 308–327.

Mullaney, J. A. (1984). The relationship between anxiety and depression: a review of some principal component analytic studies. *Journal of Affective Disorders* **7**, 139–148.

Patel, V., Gwanzura, F., Simunyu, E., Lloyd, K. & Mann, A. (1995). The phenomenology and explanatory models of common mental disorders: a study in primary care in Harare, Zimbabwe. *Psychological Medicine* **25**, 231–239.

Patel, V., Simunyu, E., Gwanzura, F., Lewis, G. & Mann, A. (1997). The Shona symptom questionnaire: the development of an indigenous measure of non-psychotic morbidity in Harare, Zimbabwe. *Acta Psychiatrica Scandinavica* **95**, 469–475.

Paykel, E. S., Prusoff, B. A. & Klerman, G. L. (1971). The endogenous neurotic continuum in depression – rater independence and factor distribution. *Journal of Psychiatric Research* **8**, 73–90.

Rosenthal, S. H. & Gudeman, J. E. (1967). The endogenous depressive pattern. *Archives of General Psychiatry* **16**, 241–249.

Tyrer, P. 1985). Neurosis divisible? *Lancet* i, 685–688.

Weich, S., Churchill, R., Lewis, G. & Mann, A. (1996). Do socio-economic factors predict the incidence and maintenance of psychiatric disorder in primary care? *Psychological Medicine* **27**, 73–80.