The factor structure of the Edinburgh Postnatal Depression scale in a South African peri-urban settlement

Gideon P. de Bruin*
Department of Psychology, University of Stellenbosch, South Africa
(Current correspondence address: Institute for Child and Adult Guidance, Rand Afrikaans University, Box 524, Auckland Park, 2006 South Africa)e-mail: gpdb@rau.ac.za

Leslie Swartz
Department of Psychology, University of Stellenbosch and Human Sciences Research Council, Cape Town, South Africa

Mark Tomlinson
Child Guidance Clinic, University of Cape Town and Winnicott Research Unit, University of Reading

Peter J. Cooper
Department of Psychology, University of Reading, United Kingdom

Christopher Molteno
Department of Psychiatry, University of Cape Town, South Africa

The factor structure of the Edinburgh Postnatal Depression scale (EPDS) and similar instruments have received little attention in the literature. The researchers set out to investigate the construct validity and reliability of the EPDS amongst impoverished South African women. The EPDS was translated into isiXhosa (using Brislin's back translation method) and administered by trained interviewers to 147 women in Khayelitsha, South Africa. Responses were subjected to maximum likelihood confirmatory factor analysis. A single factor structure was found, consistent with the theory on which the EPDS was based. Internal consistency was satisfactory (a = 0.89).

* To whom correspondence should be directed

© Psychological Society of South Africa
ISSN 0081-2463
South African Journal of Psychology 2004, 34 (1), pp. 113–121.

With increasing interest in postnatal depression worldwide, the Edinburgh Postnatal Depression scale (EPDS) is commonly used as a screening instrument (Cox, Holden & Sagovsky, 1987). Some authors assume that the EPDS is valid cross-culturally (Grace, Lee, Ballard & Herbert, 2001); others evaluate the EPDS against a 'gold standard' or criterion validity measure (Benvenuti, Ferrara, Niccolai, Valoriani & Cox, 1999; Wickberg & Hwang, 1996). In South Africa, Lawrie, Hofmeyr, de Jager and Berk (1998) validated the EPDS against the 'gold standard' of a DSM-IV diagnosis of depression (amongst women in postnatal care) at Coronation Hospital, Johannesburg. They report a sensitivity of 80% and specificity of 77% for the EPDS, very similar to the British figures (Murray & Carothers, 1990).

It is important to know whether the EPDS identifies cases correctly, but this is not the only question that should be asked about validity. Debates about the nature of depression cross-culturally have raised questions, for example, about the appropriateness of depression as a construct in non-Western contexts (Swartz, 1998), the feasibility of separating depression from anxiety and/or somatic symptoms (Mineka, Watson & Clark, 1998) and the question, in impoverished contexts, of the appropriateness of the psychopathological term *depression,* for what could be termed *demoralisation or social suffering* (cf. Kleinman & Kleinman, 1991). These questions can be addressed fully only by using multiple methods, both quantitative and qualitative. One question is whether the same construct is measured in different cultures by instruments such as the EPDS. Furthermore, (it may be the case that) in some contexts the instrument may not be mapping onto a single psychopathological category. This study explores the construct validity of the EPDS in an impoverished South African community.

The factor structure of a measuring instrument has important implications for its construct validity. Factors represent theoretical constructs that are thought to underlie or explain the covariances amongst the items that serve as indicators of the constructs. Hence, if the measurement model of an instrument specifies that a single theoretical construct is measured by the items, then a single factor should provide a satisfactory explanation of the covariances amongst the items and each item should be strongly and significantly related to the factor. Similarly, if the measurement model specifies that the items measure two theoretical constructs, then two factors should provide a satisfactory explanation of the covariances between the items. Furthermore, each item should have a significant relationship with the factor of which it is intended to be an indicator, and a weak relationship with the factor of which it is not intended to be an indicator.

METHOD

Participants

A cohort of 147 women living in Khayelitsha, an informal peri-urban settlement close to Cape Town, South Africa, were recruited as part of a larger study on the impact of

postnatal depression on infant development. Details of recruitment procedures and the study context are reported by Cooper, Tomlinson, Swartz, Woolgar, Murray and Molteno (1999). The rate of depression as measured by the Structured Clinical Interview for DSM IV (SCID-I/NP) (First, Gibbon, Spitzer & Williams, 1996) was 34.7%. For purposes of the present discussion it is sufficient to mention that we made vigorous efforts to recruit all mothers of young infants in the area.

Instruments

The ten-item EPDS was developed by Cox *et al.* (1987) as a self-report measure of postnatal depression. It has shown impressive criterion-related validity in several countries (Benvenuti *et al.*, 1999; Boyce, Stubbs & Todd, 1993; Cox *et al.*, 1987; Ghubash, Abou-Saleh & Daradkeh, 1997; Lee *et al.*, 1998; Murray & Carothers, 1990; Pop, Komproe & van Son, 1992; Wickberg & Hwang, 1996). There may, however, be difficulties in some contexts with the way items are understood by respondents. In a study conducted in India, the EPDS was translated into Hindi and the format of the scale was changed to be interviewer-administered (Banerjee, Banerjee, Kriplani & Saxena, 1999). Despite these changes, Banerjee *et al.* (1999) reported that respondents had difficulty in understanding the meaning of items that were perceived by the researchers to require introspection.

Only Pop, Komproe and van Son (1992) have addressed the factor structure of the EPDS. They performed a principal axis factor analysis on the intercorrelations of the EPDS items and extracted three factors. Subsequent confirmatory factor analyses, however, suggested that a two-factor model provided the best model-data fit. Pop *et al.* (1992) labelled the two factors *depressive feelings* and *cognitive anxiety*. In spite of this multidimensional structure, however, Pop *et al.* (1992) reported data on the criterion-related validity of the EPDS as if it were measuring one dimension. If the EPDS, as in the Pop *et al.* study, measures two or three dimensions of postnatal depression, and these dimensions are replicated in other contexts, it may become appropriate to revisit the scoring of the instrument. The question of the dimensionality of the EPDS and the implications of this, therefore, remains.

In our study, the EPDS was translated into isiXhosa (the language spoken in the community in which the study was conducted) by the translation-back translation method as described by Brislin (1986). Because of high rates of illiteracy amongst respondents, an interviewer-administered format was used (cf. Banerjee, Banerjee, Kriplani & Saxena, 1999) and respondents were asked to report on symptoms over a two-week period.

Procedure

Mothers were assessed by one of two interviewers who had been trained by the second and fourth authors of this article. The interviewers were native speakers of isiXhosa who were familiar with the community. Assessments took place at a research base in

Khayelitsha. Further details of the procedure for the study as a whole are reported elsewhere (Cooper *et al.*, 1999).

Data analysis

Maximum likelihood confirmatory factor analysis was used to test two measurement models for the EPDS. The first model, Model 1, represented the hypothesis that the EPDS measures a one-dimensional construct. Model 1 specified that scores on the EPDS items are influenced by (a) a single factor common to all the items, namely, degree of postnatal depression, and (b) a factor unique to every item. In accordance with factor theory, a unique factor reflects reliable variance that is specific to a particular variable and random measurement error (Comrey & Lee, 1992). This model is depicted in Figure 1. The metric of the latent variable or factor was established by constraining its variance to one. Additionally, the regression weights of all the items on the unique factors were constrained to be equal to one.

The second measurement model, Model 2, was based on the findings of Pop *et al.* (1992) and represented the hypothesis that the EPDS measures two correlated factors, namely, *depressive feelings* and *cognitive anxiety*. In accordance with Pop *et al*.'s findings, Items 1, 2, 6, 7, 8, 9 and 10 were specified as indicators of the *depressive feelings* factor, and Items 3, 4 and 5 were specified as indicators of *cognitive anxiety*. The correlation between the two substantive factors was freely estimated from the data. As with the first measurement model, each item was specified as being influenced by a unique factor in addition to the substantive factors. The metric of the two substantive factors was established by constraining their respective variances to one. The regression weights of the items on the unique factors were also constrained to be equal to one. This model is depicted in Figure 2.

RESULTS

The hypothesis of a perfect fit between Model 1 and the data was rejected [$\chi^2(35) = 67.44$; $p \leq 0.001$]. However, Browne and Cudeck (1992) have argued that it is unrealistic to expect a perfect fit between a hypothesised model and observed data. They suggested that the root mean square error of approximation (RMSEA) provides a more realistic index of model fit than the chi-square statistic. RMSEA values of 0.05 and lower are thought to reflect a close fit between the model and the data, and values between 0.05 and 0.08 are thought to reflect a satisfactory fit (Browne & Cudeck, 1992). Using these criteria, the value obtained in the present study (RMSEA = 0.08) can be described as marginally satisfactory. Furthermore, the values of the Goodness of Fit Index (GFI = 0.91), and the Comparative Fit Index (CFI = 0.95) all suggest a satisfactory fit between the hypothesized measurement model and the observed data. The fit indices for Model 1 are summarised in Table 1.

Table 1. Fit indices for two measurement models of the EPDS for isiXhosa-speaking women

	χ^2	*df*	GFI	CFI	RMSEA
Model 1	67.4	35	0.91	0.95	0.080
Model2	58.43	34	0.92	0.96	0.070

The standardised estimated factor pattern coefficients of the first measurement model are presented in Figure 1. These coefficients reflect the links between the observed scores on the items and the latent construct, namely postnatal depression. All factor pattern coefficients were statistically significant (r, < 0.05) and the coefficients can generally be described as moderately strong. The weakest indicators were Items 5 ($\beta = 0.39$) and 10 ($\beta = 0.42$), and the best indicators were items 6 ($\beta = 0.82$) and 9 ($\beta = 0.81$).

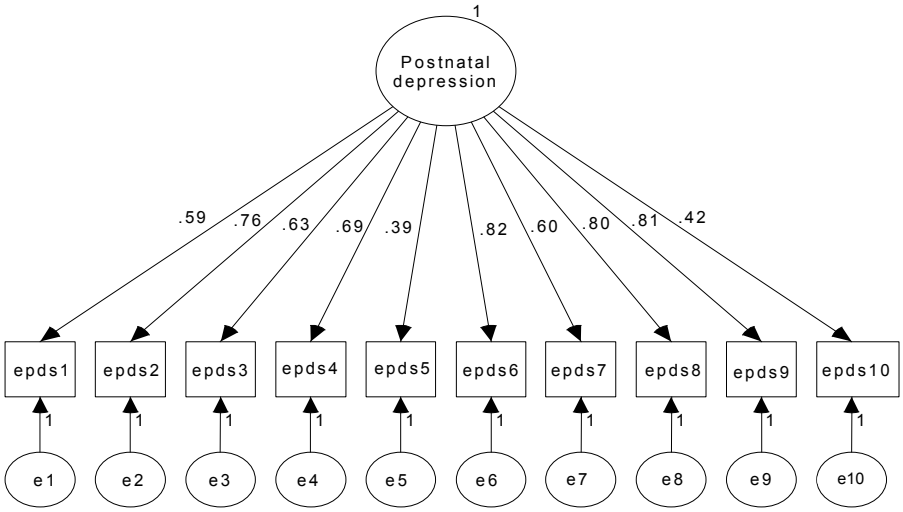

Figure 1. Standardised estimated parameters of Measurement Model 1 of the Edinburgh Postnatal Depression scale for isiXhosa-speaking women

The hypothesis of a perfect fit between Model 2 and the observed data was also rejected [χ^2 (34) = 58.43; p = 0.006]. However, the Goodness of Fit Index (GFI = 0.92), Comparative Fit Index (CFI = 0.96), and the Root Mean Square Error of Approximation (RMSEA = 0.07) all indicated a satisfactory fit between the model and

the observed data. The fit indices are summarised in Table 1. The difference between the chi-squares for the two models shows that, statistically, Model 2 fits the data significantly better [χ^2 (1) = 9.01; $p < 0.05$]. The GFI, CFI and RMSEA of Model 2, however, are only slightly better than that of Model 1 (see Table 1), suggesting that the improvement in fit is not practically meaningful.

The standardised estimated parameters for Model 2 are presented in Figure 2. All the factor pattern coefficients were statistically significant ($p < 0.05$) and relatively strong. The very strong correlation between the *depressive feelings* and *cognitive anxiety* factors are of theoretical and practical interest ($r = 0.86$; $p < 0.05$). This correlation reflects a large degree of overlap between the two factors and it appears that they basically provide the same information. In view of the observation that the one-dimensional model is (a) conceptually more simple and (b) displays only a slightly worse fit with the observed data than the two-dimensional model, it is believed that the one-dimensional model has more scientific and practical merit.

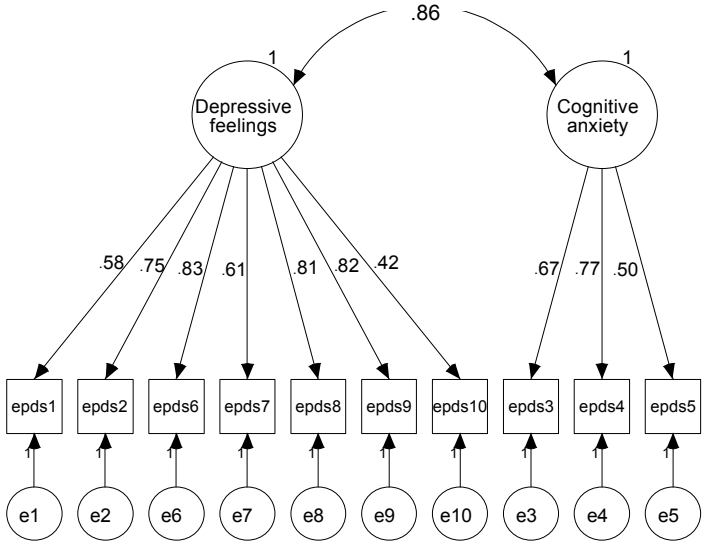

Figure 2. Standardised estimated parameters of Measurement Model 2 of the Edinburgh Postnatal Depression scale for isiXhosa-speaking women

Granted that the EPDS appears to measure an essentially one-dimensional construct, the reliability of the scores was determined. The internal consistency reliability of the total score can be described as satisfactory. The Cronbach alpha coefficient obtained ($\alpha = 0.89$) compared well with coefficients obtained in Scotland ($\alpha = 0.89$) (Cox *et al.*, 1987), the Netherlands ($\alpha = 0.80$) (Pop *et al.*, 1992), the United Arab Emirates ($\alpha = 0.84$) (Ghubash *et al.*, 1997), and Italy ($\alpha = 0.79$) (Benvenuti *et al.*, 1999). Therefore, it was concluded that reliable EPDS scores were obtained for the isiXhosa-speaking participants.

DISCUSSION

The results of the confirmatory factor analysis provide support for the construct validity of the interviewer-administered isiXhosa version of the EPDS. It appears that a single common factor underlies responses to the ten items. It is, therefore, appropriate to sum the ten items to obtain a total score. As we have shown, the two-factor model is not superior to the one-factor model for this group of participants. What counts in favour of the present model is that it is parsimonious and consistent with the theory on which the EPDS is based. The internal consistency analysis further indicated that reliable scores were obtained with the isiXhosa version of the EPDS. These results are encouraging and suggest that there is value in pursuing the question of the use of the EPDS as a screen for postnatal depression with isiXhosa-speaking South African women.

All items loaded satisfactorily on the single construct, using the convention that 0.30 represents a satisfactory factor loading. It is, however, interesting to note that the two items which loaded relatively weakly on the common factor in this study were Items 5 ('Have you been feeling scared or panicky for no good reason?') and 10 ('Have things been so bad that the thought of harming yourself occurred to you?'). Item 5 can reasonably be considered a measure of anxiety or fear rather than depression *per se*. Given the violent and insecure nature of life in Khayelitsha, one might have expected that substantial numbers of women would have responded in the affirmative to this item, but this was not in fact the case, with 91.2% of respondents not scoring on this item. The vast majority of women, furthermore, did not report any thoughts of self-harm (85.7%).

The results of the confirmatory factor analysis suggest that the symptoms of postnatal depression among isiXhosa women manifest in a way that is consistent with theoretical descriptions and observations obtained from Europe and North America. Though broader debates about the nature of depression cross-culturally remain, it is possible that isiXhosa women and European women may suffer from a similar affective disorder after childbirth. The challenge now is to link ethnographic and psychometric studies to gain further insight into the nature and context of postnatal depression in the developing world.

ACKNOWLEDGEMENTS

We are grateful to the mothers who participated in the study. Thanks are due to Nokwanda Mtoto for assistance with participant recruitment, to Nosisana Nama and Nomalanga Mosala for carrying out the assessments, and to Robin Sandler for help with data entry. The initiation of the Winnicott Research Unit–South Africa collaboration was made possible by a British Council Link Award. This study was supported by a grant from the Wellcome Trust of Great Britain.

REFERENCES

Banerjee, N., Banerjee, A., Kriplani, A. and Saxena, S. (1999). Evaluation of postpartum depression using the Edinburgh Postnatal Depression scale in evaluation of postpartum depression in a rural community in India. *International Journal of Social Psychiatry, 46*, 74-75.

Benvenuti, P., Ferrara, M., Niccolai, C., Valoriani,V. and Cox, J. L. (1999). The Edinburgh Postnatal Depression scale: Validation for an Italian sample. *Journal of Affective Disorders, 53*, 137-141.

Boyce, P., Stubbs, J. and Todd, A. (1993). The Edinburgh Postnatal Depression scale: Validation for an Australian sample. *Australian and New Zealand Journal of Psychiatry, 27*, 472-476.

Brislin, R.W. (1986). The wording and translation of research instruments. In W.J. Lonner and J.W. Berry (Eds), *Field methods in cross-cultural research* (pp. 137-164). Newbury Park, CA: Sage.

Browne, M.W. and Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*, 230-258.

Comrey, A.L. and Lee, H.B. (1992). *A first course in factor analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Cooper, P.J., Tomlinson, M., Swartz, L., Woolgar, M., Murray, L. and Molteno, C. (1999). Postpartum depression and the mother–infant relationship in a South African peri-urban settlement. *British Journal of Psychiatry, 175*, 554-558

Cox, J.L., Holden, J.M. and Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression scale. *British Journal of Psychiatry, 150*¸ 782-786.

First, M.B., Gibbon, M., Spitzer, R.L and Williams, J.B.W. (1996). *Structured cinical iterview for DSM-IV axis I.* New York: New York State Psychiatric Institute.

Ghubash, R., Abou-Saleh, M.T. and Daradkeh, T.K. (1997). The validity of the Arabic Edinburgh Postnatal Depression scale. *Social Psychiatry and Psychiatric Epidemiology, 32*¸ 474-476.

Grace, J., Lee, K.K., Ballard, C. and Herbert, M. (2001). The relationship between postnatal depression, somatization and behaviour in Malaysian women. *Transcultural Psychiatry, 38*, 27-34.

Kleinman, A. and Kleinman, J. (1991). Suffering and its professional transformation: Towards an ethnography of interpersonal experience. *Culture, Medicine and Psychiatry, 15*¸ 275-301.

Lawrie, T.A., Hofmeyr, G.J., de Jager, M. and Berk, M. (1998). Validation of the Edinburgh Postnatal Depression scale on a cohort of South African women. *South African Medical Journal, 88*, 1340-1344.

Lee, D.T.S., Yip, S.K., Chiu, H.F.K., Leung, T.Y.S., Chan, K.P.M., Chau, I.O.L., Leung, H.C.M. and Chung, T.K.H. (1998). Detecting postnatal depression in Chinese women. *British Journal of Psychiatry, 172*, 433-437.

Mineka, S., Watson, D. and Clark, L.A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377-412.

Murray, L. and Carothers, A.D. (1990). The validation of the Edinburgh Post-Natal Depression scale on a community sample. *British Journal of Psychiatry*, *157*, 288-290.

Pop, V.J., Komproe, I.H. and van Son, M.J. (1992). Characteristics of the Edinburgh Post-Natal Depression scale in the Netherlands. *Journal of Affective Disorders, 26,* 105-110.

Swartz, L. (1998). *Culture and mental health: A Southern African view.* Cape Town: Oxford.

Wickberg, B. and Hwang, C.P. (1996). The Edinburgh Postnatal Depression scale: Validation on a Swedish community sample. *Acta Psyciatrica Scandinavia, 94*, 181-184.