Routledge
Taylor & Francis Group

Journal of Reproductive and Infant Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjri20

Multidimensionality within the Edinburgh postnatal depression scale: application issues of specific structure

Kateřina Ratislavová, Eva Hendrych Lorenzová, Alena Lochmannová & Colin R. Martin

To cite this article: Kateřina Ratislavová, Eva Hendrych Lorenzová, Alena Lochmannová & Colin R. Martin (23 Nov 2023): Multidimensionality within the Edinburgh postnatal depression scale: application issues of specific structure, Journal of Reproductive and Infant Psychology, DOI: 10.1080/02646838.2023.2285837

To link to this article: https://doi.org/10.1080/02646838.2023.2285837

Published online: 23 Nov 2023.

JOURNAL OF REPRODUCTIVE AND INFANT PSYCHOLOGY
https://doi.org/10.1080/02646838.2023.2285837

Multidimensionality within the Edinburgh postnatal depression scale: application issues of specific structure

Kateřina Ratislavová[a], Eva Hendrych Lorenzová[a], Alena Lochmannová[b] and Colin R. Martin[c]

[a]Midwifery, Department of Nursing and Midwifery, Faculty of Health Care Studies, University of West Bohemia, Pilsen, Czech Republic; [b]Department of Paramedic Science, Medical Diagnostics and Public Health, Faculty of Health Care Studies, University of West Bohemia, Pilsen, Czech Republic; [c]Professor of Clinical Psychobiology and Applied Psychoneuroimmunology, Institute for Health and Wellbeing, University of Suffolk, Ipswich, UK

ABSTRACT

Objective and background: The 10-item Edinburgh Postnatal Depression Scale (EPDS) is a widely-used screening measure for postnatal depression. Factor analysis studies have suggested an embedded sub-scale could be used for screening for anxiety disorders. The current investigation sought to replicate and extend a recent study supporting this assertion.
Methods: A cross-sectional design. EPDS data were collected at up to two years postpartum. Confirmatory factor analysis, correlational and distributional characteristics of the measure were examined. Participants were a large sample ($N = 985$) of postpartum women in the Czech Republic.
Results: Factor structure findings substantially replicated the models evaluated by Della Vedova et al. (2022). Bifactor models, however, offered a better fit to data. A general factor of depression explained most of the variance in data in most models compared to embedded sub-scales across models.
Conclusion: The model proposed by Della Vedova et al. (2022) offered an excellent fit to data. However, the findings from the bifactor modelling suggest the dominance of a general factor of depression which indicates the potential application of an embedded anxiety sub-scale for screening may be overstated.

ARTICLE HISTORY
Received 17 January 2023
Accepted 15 November 2023

KEYWORDS
Edinburgh Postnatal Depression Scale (EPDS); factor structure; validity; psychometrics; screening

Introduction

Postnatal depression (PND) affects a significant minority of new mothers (O'Hara & Swain, 1996) and impacts deleteriously not only on the mental health of the mother herself, but also the infant and partner (Cameron et al., 2017; Dahlen et al., 2015; Giallo et al., 2015; Pope et al., 2013). Screening guidelines for PND vary widely (American College of Obstetricians and Gynecologists' Committee on Obstetric Practice, 2015; National Institute for Health and Care Excellence, 2015; Scottish Intercollegiate Guidelines

CONTACT Colin R. Martin C.Martin6@uos.ac.uk Institute for Health and Wellbeing, University of Suffolk, Ipswich IP4 1QJ, UK

Network, 2012). One step and two step (follow-up to positive response to a minimal screen) are frequently recommended. The most commonly utilised screening tool for PND is the Edinburgh Postnatal Depression Scale (EPDS (Cox et al., 1987)), the instrument itself being globally endorsed as a clinical outcome measure of choice (The International Consortium for Health Outcome Measurement, 2017). The tenure of the EPDS as the 'gold standard' screening measure for PND shows no sign of abating, however, conjecture regarding whether the tool does indeed measure a single dimension of depression (as originally conceived and currently operationalised) has been a persistent area of concern and indeed interest, within the PND screening literature. Accruing evidence suggests the EPDS may indeed be a multi-dimensional measure (Brouwers et al., 2001; Jomeen & Martin, 2005, 2007; Kozinszky et al., 2017; Matthey, 2008; Phillips et al., 2009; Reichenheim et al., 2011). The implications of multi-dimensionality within the EPDS are not only for screening veracity in terms of the underlying tenet of measuring a unidimensional construct (depression) but also in relation to potentially useful sub-scales which may be embedded within the tool that could screen for anxiety and anhedonia (Jomeen & Martin, 2005; Kozinszky et al., 2017; Matthey et al., 2013). A sense of therapeutic optimism has been garnered by the potential use of embedded sub-scales identified following the original development of the measure (Kozinszky et al., 2017; Phillips et al., 2009). In particular, the potential to use the EPDS to *potentially* screen for both anxiety and depression has long-standing and contemporary support (Della Vedova et al., 2022; Gollan et al., 2017; Matthey et al., 2013). Martin and Redshaw (2018) confirmed that when evaluating various multi-dimensional models of the EPDS at different time-points postpartum, the measurement models remained stable. Thus, the conceptual underpinnings of the models evaluated in that study exhibited measurement coherence in contrast to measurement transience and so furnished additional evidence for the potential application of the EPDS as a sub-scaled screening measure. Interestingly, scrutiny of individual EPDS items suggests, from a face validity point of view, that two of the items might be more readily considered anxiety items, for example, item 4. 'I have been anxious or worried' and item 5. 'I have felt scared or panicky', though contemporary psychological models of depression consider anxiety elements to be a component of depression (Clark & Watson, 1991). Under this rubric the application of the EPDS as a unitary measure of depression does not represent a theoretical misspecification, however, the consistent findings from factor analysis studies suggest the presence of multi-dimensionality does indicate a measurement model misspecification. Thoughthese seemingly opposing positions have not to date been reconciled with any degree of confidence, the presence of multidimensionality does in principle, albeit serendipitously, support the potential to use the measure as a sub-scaled tool. Indeed, the gravitas accorded to factor analysis studies indicating an anxiety sub-scale within the EPDS may be inferred by the inclusion of EPDS items 3,4 and 5 as a potential screening measure for perinatal anxiety disorders in at least one national perinatal mental health clinical guideline (Austin et al., 2017).

A recent study undertaken in Italy (Della Vedova et al., 2022) using an elegant methodology of exploratory factor analysis (EFA) followed by confirmatory factor analysis (CFA) on a suitably large sample indicated a two-factor (anxiety and depression) structure to the EPDS and moreover the suggested application of these two embedded sub-scales for screening purposes. The anxiety sub-scale, the EPDS-4A comprised items 3,4,5 and 6. It is noteworthy that this represents an extension to the proposed EPDS-3A that comprises items 3,4 and 5 for anxiety screening (Matthey, 2008), an EPDS sub-scale formulation that has also found empirical support recently by Smith-Nielsen et al. (2021) suggesting the veracity of this sub-scale for antenatal and postnatal screening. Closer scrutiny of the recent Italian study (Della Vedova et al., 2022) does highlight some equivocal observations. Firstly, it was noted that the EFA indicated a split-loading item (or perhaps more accurately, a non-loading item < 0.30) on anxiety and depression factors in relation to item 6. Traditionally, an EFA approach would be to reject such an item (Kline, 1994). Della Vedova et al. (2022) recognised this problem and thus run three alternative CFA models, the EPDS-4A model above, a similar model but with item-6 loading on the depression factor, and a two-factor 9-item model excluding item-6. Scrutiny of all three of these CFA models does reveal that none provide a satisfactory fit to data. It should also be noted that the model with item-6 loading on depression has previously been highlighted as an optimal factor analysis solution and interpretation of EPDS data by Gollan et al. (2017). Attempting to address this issue, the authors examined the model modification indices and elected to correlate error terms between item-5 and items 7 and 10. This approach offered a good fit to the data and was further supported by an unambiguously excellent fit to data of the four items of the EPDS-4A. Correlation of error terms within a CFA model is circumscribed only on the basis of robust theoretical grounds (Byrne, 2010), a perspective the authors recognised in the justification for their approach. However, correlation of error terms *between* rather than *within* factors is discouraged as an erroneous practice within the CFA and structural equation modelling literature (Byrne, 2010). Under this rubric, Della Vedova et al. (2022) may have encountered a methodological conundrum, namely, a good model fit to data as a consequence of oblique statistical propriety with respect to correlating errors across factors. A further issue concerns the notion of the potential superiority of a bifactor model in explaining model fit to data (Ratislavova et al., 2022). Essentially, a bifactor model establishes within a structural model the variance that is explained by a single global factor, *in contrast*, to that explained by sub-scales purported to be implicit to the measure itself. In effect, the bifactor model is extremely useful in identifying if sub-scales are substantive within a measure once the variance explained by the global factor is accounted for (Martin et al., 2018). The approach has found merit in measures where, for example, the instrument may be used as both a total scored instrument and a sub-scaled tool. The Birth Satisfaction Scale-Revised (Hollins Martin & Martin, 2014), which can be used in total score and sub-scale score guises depending on research or clinical purposes, has been found to be supported in both approaches by the use of bifactor modelling (Martin et al., 2018). Interestingly, a bifactor model was found to offer a superior fit to the EPDS by Reichenheim et al. (2011). This methodologically complex study did indeed find support for a general factor of depression and also some support for three defined factors within the instrument. However, the ambiguous fidelity of the three factors identified within the data suggested to the researchers that these factor-derived sub-scales within the EPDS should not be used in practice (Reichenheim et al., 2011). A study by Martin and Redshaw (2018) was unable to replicate this specific bifactor model since data fit to model yielded an unanalysable solution.

The objectives of the current study are to:

(1) Replicate the measurement model evaluations of the EPDS undertaken by Della Vedova et al. (2022) in a large sample of Czech postpartum women.
(2) Evaluate competing best-fit models of the EPDS identified by Martin and Redshaw (2018) with those recently reported by Della Vedova et al. (2022).
(3) Evaluate alternative bifactor models of the EPDS including those proposed Della Vedova et al. (2022).
(4) Evaluate the pragmatic reasonableness of differentiation of hypothesised sub-scales for applied use by examination of the relationship of such sub-scales to related constructs.

Methods

Design

A cross-sectional design was used for the study. Inclusion criteria included speaking Czech, age 18 years or older, having given birth within the past 24 months. The participants were recruited by means of an online survey using convenience sampling. Informed consent for study participation was embedded in the survey. The call for participation and the online questionnaire was published in six different forums for women on maternity and parental leave in the Czech Republic.

To address objectives 1 and 2, the unidimensional and multi-dimensional models reported by Della Vedova et al. (2022) and Martin and Redshaw (2018) were evaluated using CFA. Due to the large number of models evaluated by Martin and Redshaw (2018) in the interests of parsimony and potential clinical application, we selected only those models which used either nine or the full ten items. Shortened versions such as the six-item model of Kozinszky et al. (2017) which though theory-driven, have not to date been applied in clinical screening practice and thus these shortened measures were not included if more than 1-item was removed. To address objective 3. bifactor modelling was used which represents a specific type of structural equation model (SEM), distinct from CFA and indeed it should be recognised that CFA is a special case of a SEM. To address objective 4. best-fit models were correlated with anticipated related constructs including birth experience and posttraumatic stress disorder (PTSD) to determine variation in the relationship to these constructs in order to establish the degree of sub-scale differentiation from the core concept of depression. The 10-item total EPDS score was also used in this analysis as a reference.

Ethical approval

Ethical approval was gained from The Research Ethics Committee of the University of West Bohemia in Pilsen, reference number ZCU 000065/2023. The Research Ethics Committee's decision confirmed that the study participants are volunteers, their human dignity is not violated, and they have not been exposed to physical, psychological, or social risks. Only anonymous data were collected from the study participants and their privacy and personal data protection is guaranteed according to the relevant law.

Measures

The Edinburgh Postnatal Depression Scale (EPDS)

The EPDS (Cox et al., 1987) is a widely-used and guideline recommended screening measure of PND comprising tenself-report items and scored on 4-point Likert-like endorsement scale (range = 0–30). The recommended 'cut-score' threshold is 12/13 for probable PND (Cox et al., 1987) with a lower screening threshold of 9/10 also suggested (Cox & Holden, 2003). A Czech language validated-version of the EPDS (Břicháček et al., 2000) was used in the current study and we used the optimal threshold score for this version of 10/11 determined by the sensitivity and specificity analysis of Horakova et al. (2022).

The Birth satisfaction Scale-Revised (BSS-R)

The Birth Satisfaction Scale-Revised (BSS-R (Hollins Martin & Martin, 2014)), is a validated 10-item multidimensional self-report measure of birth experience. The BSS-R can be used in total score and sub-scale forms depending on application (Martin et al., 2018). The BSS-R is recommended for global use (The International Consortium for Health Outcome Measurement, 2017) and has recently been translated and validated in Czech (Ratislavova et al., 2022). Birth experience has been found to be associated with postnatal depression (Bell & Andersson, 2016) and specific dimensions of the BSS-R have been found to predict EPDS scores (Nakic Rados et al., 2022)

The Primary Care-Posttraumatic Stress Disorder (PC-PTSD-5) scale

The Primary Care-Posttraumatic Stress Disorder scale (PC-PTSD-5) (Prins et al., 2015), is a 5-item self-report measure designed to screen for PTSD within a primary care setting. We modified the instrument for use within the present study by changing the stem to ask how traumatic the woman perceived the most recent birth. The 5-items are then contextualised to the birth in this modification before translation into Czech. Cut-point screening recommendations for the PC-PTSD-5 vary between 3-4 of the five questions positively endorsed (Prins et al., 2015). Significant relationships have been observed between EPDS scores, reports of a traumatic birth experience and postpartum PTSD (Nagle et al., 2022; Yakupova & Suarez, 2022). We elected to adapt the PC-PTSD-5 for our study based on linkage to DSM-V criteria, brevity and the initial pilot testing of our study questionnaires.

Statistical analysis

SEM and CFA were used to address all research objectives (Brown, 2015). It is noted that Martin and Redshaw (2018) used an alternative estimation method to maximum-likelihood in their study due to the EPDS item 10. 'The thought of harming myself has occurred to me' being observed to be highly skewed and exceptionally kurtotic. This may have been a particular population-bound characteristic of their sample, thus in the current study skew and kurtosis of each item will be examined to determine the

Table 1. Theoretically and factor-derived models with item-factor relationships indicated.

Model	N Factors	Factor 1 (F.1)	F.1 items	Factor 2 (F.2)	F.2 items	Factor 3 (F.3)	F.3 items
Cox et al. (1987)	1	Depression	1–10				
Della Vedova et al. (2022)	2	Depression	1–2, 7–10	Anxiety	3–6		
Gollan et al. (2017) two-factor*	2	Depression	1–2, 6–10	Anxiety	3–5		
Della Vedova et al. (2022) nine-item	2	Depression	1–2, 7–10	Anxiety	3–5		
Reichenheim et al. (2011) Three-factor	3	Depression	7–10	Anxiety	3–5	Anhedonia	1–2, 6
Tuohy and McVey (2008) Three-factor	3	Depression	7–10	Anxiety	1–2	Anhedonia	3–5
Coates et al. (2016) Three-factor	3	Depression	7–10	Anxiety	3–6	Anhedonia	1–2

*This model is identical to that of the modified model of Della Vedova et al. (2022) with item-6 loading on depression. Since Gollan and colleagues study predates that of Della Vedova et al. (2022), it would seem more appropriate to describe this model as Gollan et al. (2017).

appropriateness of maximum-likelihood estimation or whether a correction or alternative estimation approach may be required.

The CFA models are summarised in Table 1, with the salient features described including number of factors, item-factor loadings, and factor-content description.

Consistent with the approach taken by Martin and Redshaw (2018), models were evaluated using a variety of model fit indices (Bentler & Bonett, 1980). These were, the comparative fit index (CFI) (Bentler, 1990); the root mean squared error of approximation (RMSEA) (Steiger & Lind, 1980) and the squared root mean residual (SRMR) (Hu & Bentler, 1999). CFI values > 0.95 indicates a good fit to the data (Hu & Bentler, 1999). RMSEA values ≤ 0.08 indicate acceptable model fit (Browne & Cudeck, 1993), and values of ≤ 0.05 indicative of good fit (Schumacker & Lomax, 2010). SRMR values of < 0.08 are indicative of good model fit (Hu & Bentler, 1999). The χ^2 is generally not used for model evaluation due to sample size inflation effects (Byrne, 2010). The same fit indices and criteria are applied to bifactor models, however in terms of the modelling the specific factors are specified as uncorrelated and the relationship between domain-specific factors and the general factor also being specified as orthogonal. This approach is used in order to allow the additional variance contribution of domain-specific factors to be determined over and above that of the general factor variance explained (Martin et al., 2018).

The relationship between EPDS total and model circumscribed sub-scales and the BSS-R and the PC-PTSD-5 was undertaken using Pearson's *r* correlation coefficient. Comparisons between the total EPDS score and those reported by Martin and Redshaw (2018) were undertaken using the one-sample *t*-test.

Statistical analysis was undertaken using the R programming language (v.4.2.2) (R Core Team, 2022), and the R packages Lavaan (Rosseel, 2012), semTools (Jorgensen et al., 2022), Psych (Revelle, 2022), Psy (Falissard, 2022) Hmisc (Harrell, 2022) and QuantPsyc (Fletcher, 2022) were used for all statistical analyses.

Results

EPDS data from one-thousand and seventeen women were collected. It was noted that three participants did not complete item 9 of the BSS-R and one participant did not

Table 2. Characteristics of the sample.

Variable	N	Mean	SD	Min	Max	Kurtosis	Skew	S. error
Age (years)		30.76	4.59	18	44	0.14	−0.1	0.15
Months since birth		12.49	7.32	0	24	−0.03	−1.2	0.23
BSS-R total score		23.14	8.99	0	40	−0.20	−0.88	0.29
BSS-R SE sub-scale		8.10	4.32	0	16	−0.01	−1.01	0.14
BSS-R WA sub-scale		4.09	2.40	0	8	−0.15	−1.12	0.08
BSS-R QC sub-scale		10.96	3.73	0	16	−0.51	−0.67	0.12
PC-PTSD total score		1.39	1.44	0	5	0.92	−0.04	0.05
Birth type (spontaneous vaginal/induced vaginal/emergency CS/planned CS/forceps or ventouse)	426/259/162/62/76 (43/26/17/6/8%)							
Parity (null/multip)	648/337 (66/34%)							
Hospital born (yes/no)	966/19 (98/2%)							
Term pregnancy (pre-term/term/post-term)	95/825/65 (9/84/7%)							

Table 3. Distributional characteristics of individual EPDS items and the total score ($N = 985$). SE = standard error.

Item	Mean	SD	Min	Max	Skew	Kurtosis	SE
EPDS1	0.77	0.89	0	3	0.87	−0.23	0.03
EPDS2	0.61	0.83	0	3	1.30	0.91	0.03
EPDS3	1.39	0.99	0	3	0.00	−1.06	0.03
EPDS4	1.23	0.93	0	3	0.18	−0.93	0.03
EPDS5	0.79	0.91	0	3	0.82	−0.46	0.03
EPDS6	1.34	0.86	0	3	−0.03	−0.78	0.03
EPDS7	0.99	1.04	0	3	0.62	−0.89	0.03
EPDS8	1.31	0.97	0	3	0.16	−0.99	0.03
EPDS9	0.77	0.87	0	3	0.93	0.03	0.03
EPDS10	0.23	0.60	0	3	2.86	7.81	0.02
Total	9.43	6.60	0	30	0.61	−0.20	0.21

complete item 1 of the EPDS and these cases were excluded, thus complete data from 1013 participants was subject to further scrutiny. Complete EPDS, BSS-R and PC-PTSD data were screened for multivariate outliers by reference to Mahalanobis distance from the centroid (Mahalanobis, 1936) and 20 EPDS, 10 BSS-R and 5 PC-PTSD scale outliers were identified and excluded, a small number of outliers ($N = 8$) were common to two or more of the instruments thus the final dataset comprised 985 participants. The characteristics of the sample and BSS-R and PC-PTSD scores are shown in Table 2.

To facilitate comparison with Martin and Redshaw's (2018) study, the means, standard deviations, skew and kurtosis of the EPDS data are summarised in Table 3. The mean EPDS score in the current study was observed to be significantly higher than those reported by Martin and Redshaw (2018) at three-month postpartum, $t(984) = 14.50$, $p < 0.001$, and at six-month postpartum, $t(984) = 9.12$, $p < 0.001$, means reported by Martin and Redshaw (2018) at three and six months being 6.38 and 7.51, respectively.

Confirmatory factor analysis

CFA evaluation of each model is summarised in Table 4. The bifactor models were observed to offer a (marginally) better fit to the data compared to the correlated measurement models. Specifically, the bifactor models of Della Vedova et al. (2022)

Table 4. Confirmatory factor analysis of EPDS models.

Model	N-items	χ^2	df	CFI	RMSEA	SRMR
Cox et al. (1987)	10	336.568	35	0.938	0.094	0.041
Della Vedova et al. (2022)	10	255.081	34	0.954	0.081	0.035
Gollan et al. (2017) two-factor	10	282.579	34	0.949	0.086	0.038
Della Vedova et al. (2022) nine-item	9	222.816	26	0.952	0.088	0.036
Reichenheim et al. (2011) Three-factor	10	271.613	32	0.950	0.087	0.038
Tuohy and McVey (2008) Three-factor	9	86.804	24	0.985	0.052	0.024
Coates et al. (2016) Three-factor	10	108.423	32	0.984	0.049	0.024
Bifactor model of Della Vedova et al. (2022)	10	65.196	25	0.992	0.040	0.017
Bifactor Gollan et al. (2017) two-factor	10	67.941	25	0.991	0.042	0.017
Bifactor nine-item model of Della Vedova et al. (2022)	9	53.572	18	0.991	0.045	0.017
Bifactor model of Coates et al. (2016)	10	78.534	26	0.989	0.045	0.021
Bifactor model of Reichenheim et al. (2011)	10	92.774	25	0.986	0.052	0.024
Bifactor model of Tuohy and McVey (2008)	9	61.856	19	0.990	0.048	0.021

Table 5. Correlations of EPDS total score and sub-scales by item with the BSS-R and PC-PTSD.

	PC-PTSD	1–10	1–2, 7–10	3–6	1–2, 6–10	3–5	7–10	1–2, 6	1–2
BSS-R	−0.52	−0.37	−0.36	−0.34	−0.37	−0.32	−0.36	−0.33	−0.28
PC-PTSD		0.61	0.57	0.58	0.59	0.56	0.56	0.54	0.48
1–10			0.96	0.92	0.97	0.89	0.93	0.90	0.81
1–2, 7–10				0.77	0.99	0.73	0.96	0.90	0.87
3–6					0.82	0.98	0.76	0.77	0.62
1–2, 6–10						0.76	0.95	0.93	0.85
3–5							0.73	0.69	0.59
7–10								0.77	0.69
1–2, 6									0.94

All correlations are statistically significant ($p < 0.01$).

and the two-factor model of Gollan et al. (2017) offered the best fit to data. The bifactor model of the three-factor model proposed by Coates et al. (2016) was also found to offer an excellent fit to data as did the bifactor model of Reichenheim et al. (2011). Scrutiny of these specific bifactor models in terms of item-factor loadings reveals that most of the variation in data is explained by the general factor. The three-factor correlated models of Coates et al. (2016) and Tuohy and McVey (2008) were also observed to offer an excellent fit to data. The unidimensional model of Cox et al. (1987) was observed to offer a good fit to data across most fit indices though comparatively, had the least adequate fit of all models evaluated.

Comparison with BSS-R and PC-PTSD-5

Correlations between sub-scales based on Della Vedova et al. (2022), Coates et al. (2016), Reichenheim et al. (2011), Tuohy and McVey (2008), the EPDS measurement model of Cox et al. (1987) and BSS-R and PC-PTSD scores are shown in Table 5. Correlations between the BSS-R were greater between the EPDS total score and the depression subscale of Gollan et al. (2017) compared to all other EPDS-derived sub-scales. Correlations between the PC-PTSD-5 were greater between the EPDS total score and all other EPDS-derived sub-scales.

Discussion

The findings from the current study provide a thought-provoking insight into not only the issue of multidimensionality within the EPDS, but also the practical application of assumed embedded sub-scales within the measure. Prior to an examination of the dimensionality of the EPDS in the current study, an appraisal of the distributional characteristics of the measure and in particular, comparison with the study of Martin and Redshaw (2018) is warranted. It was noted that the mean EPDS score in the current study was significantly higher than that observed in both observation points in the UK study. Moreover, examination of the item skew and kurtosis statistics illuminates an intrinsic difference in profile in relation to item-10 between the two studies. Item 10 was substantially less skewed and kurtotic compared to Martin and Redshaw (2018) suggesting comparatively more endorsement of this item. This is of particular interest because of the notion of the differential use of this particular item as a suicide screener (Howard et al., 2011). An excellent recent systematic review of suicidality measures has raised the issue of the veracity of item 10 of the EPDS in this guise, particularly in terms of potential ambiguity in the interpretation of this question by women (Dudeney et al., 2023). Compared to Martin and Redshaw (2018) our findings in relation to distributional characteristics would suggest this item is more commonly endorsed thus the profile of depression symptomology may be different in this population, specifically in relation to the women's response to the questionnaire. This is an important observation not only with respect to understanding the complexities of depression symptoms in the postnatal period but also in relation to the impact on the underlying measurement model of the instrument which will be discussed later in relation to the factor analysis findings.

Bifactor models were found to offer a better fit to data compared to correlated models. Though it is recognised that the statistical approach to bifactor modelling may confer a bias in terms of better fit to data (Morgan et al., 2015), it is noteworthy that examination of the general factor of these models reveals that little unique variance was attributable to the specific hypothesised sub-scales of the model's evaluation once the general factor variance had been accounted for. This observation contrasts markedly with the findings of bifactor analysis of other measures, for example, the BSS-R (Martin et al., 2018) where the bifactor model offers a good fit to data but the specific sub-scale factors can still be differentiated in terms of variance apportion. Consequently, in contrast to Della Vedova et al. (2022), the findings regarding the CFA's would suggest the use of the EPDS as total score unidimensional measure in terms of screening practice. Further support is garnered for this assertion by the correlations between the EPDS total score, EPDS-derived sub-scales and BSS-R and PC-PTSD-5 scores. It was observed that the correlations were greater between these two constructs and the total EPDS score than the EPDS-derived sub-scales with the sole exception of the depression sub-scale of Gollan et al. (2017) which was identical to that of the EPDS total score in terms of BSS-R correlation. Thus, given that birth experience and perceived birth trauma are established to be related to not only depression, but anxiety also (Malouf et al., 2022; Sakalidis et al., 2022), the observation that the anxiety and anhedonia sub-scales from the multidimensional models had less association with BSS-R and PC-PTSD-5 scores compared to the total EPDS score would also indicate a fundamental issue regarding the *clinical* application of such sub-scales of screening purposes, particularly for anxiety.

Taken in the round, our findings are considered with regard to the theoretical observations, thus it can be concluded that the EPDS is multidimensional, in contrast to the clinical applications, essentially, that in practical clinical use, the measure might be considered and applied within a unidimensional assumptive paradigm for screening for depression only. The bifactor models offered better fit to the data than correlation multidimensional models but within each model type (bifactor vs. correlated), there was little difference between the best fit models. It is entirely plausible to assume that differences between populations as a function of birth type, culture, parity and indeed the translation of the instrument may lead to trivial variation in factor structure which may account for the similarity between models evaluated, for example, the models of Coates et al. (2016), Tuohy and McVey (2008) and Reichenheim et al. (2011) all share an identical depression sub-scale. The utility of embedded sub-scales for screening for anxiety and comprising few items, may have both limited utility and limited appeal if population-specific characteristics may impact on the differentiation of factors across populations in a commonly used screening tool. Tempering theoretical relevance, which undoubtedly the recent work of Della Vedova et al. (2022) contributes to the literature, with the practical implications of application of their findings, the conclusion from the current study is that the use of the EPDS-4A or indeed other EPDS-derived anxiety screening sub-scales is likely to be over-stated in terms of a recommendation for clinical practice. It is not difficult to see how these sub-scale derivations of the original instrument may occur based on previous factor analysis studies, it is reasonable and evidence-based, but the findings from the current study present novel evidence of a unitary construct of depression of the tool. Thus, in light of this, the appropriateness of these sub-scales for clinical screening practice could, or perhaps should, be reviewed. Finally, we have reflected carefully on the large number of factor analysis studies that have circumscribed our current investigation to consider carefully our contribution. Firstly, in contrast to most previous EPDS factor analysis studies we have evaluated each model with an alternative bifactor modelling approach to consider the issue of unidimensionality of the measurement model in a fundamentally different way, one that extracts the uniqueness of hypothesised sub-scales within a model with precision in order to facilitate a determination to sub-scale significance in both theoretical content and statistical magnitude. This approach represents a stepwise departure to the vast majority of previous studies. Secondly, our study has attempted to leverage our findings within the applied context of the application of embedded scales that have been proposed by previous studies, thus to elaborate from studies that have been statistical and/or theoretical in nature to the practical considerations of application to screening practice. A limitation of our study was that the participant group were up to two-year postpartum. It would be useful in a future replication of the current study to undertake this type of analysis at the point of a typical postpartum depression screen. A further limitation of the study was the use of an adapted version of the PC-PTSD-5 which has yet to be validated in this population. During pilot testing of our questionnaires for the study, we noted that some women reported an issue with the measure we had originally considered using. We therefore elected to adapt the PC-PTSD-5 for the current study but we recognise that the lack of psychometric validation of the tool in this population is a limitation that we would seek to address in future research.

Acknowledgments

We are very grateful to all the participants who took part in our study and also for the expert and very helpful comments by two anonymous reviewers on a previous version of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Part of this research was supported by a foreign scientist award to Professor Martin by the University of West Bohemia, Czech Republic.

References

American College of Obstetricians and Gynecologists' Committee on Obstetric Practice. (2015). The American College of Obstetricians and gynecologists committee opinion no. 630. Screening for perinatal depression. *Obstetrics and Gynecology*, *125*(5), 1268–1271. https://doi.org/10.1097/01.AOG.0000465192.34779.dc

Austin, M.-P., Highet, N., & The Expert Working Group. (2017). *Mental Health Care in the Perinatal Period: Australian Clinical Practice Guideline*. Melbourne: Centre of Perinatal Excellence.

Bell, A. F., & Andersson, E. (2016). The birth experience and women's postnatal depression: A systematic review. *Midwifery*, *39*, 112–123. https://doi.org/10.1016/j.midw.2016.04.014

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. http://www.ncbi.nlm.nih.gov/pubmed/2320703

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Břicháček, V., Břicháčková, V., & Urbanová, M. (2000). *Edinburghská škála postnatální deprese : ověření v České republice*. Advance online publication.: https://www.medvik.cz/link/MED00113619

Brouwers, E. P., van Baar, A. L., & Pop, V. J. (2001). Does the Edinburgh postnatal depression scale measure anxiety? *Journal of Psychosomatic Research*, *51*(5), 659–663. https://www.ncbi.nlm.nih.gov/pubmed/11728506

Brown, T. (2015). *Confirmatory factor analysis for applied Research* (2nd ed.). Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2nd ed.). Routledge/Taylor and Francis Group.

Cameron, E. E., Hunter, D., Sedov, I. D., & Tomfohr-Madsen, L. M. (2017). What do dads want? Treatment preferences for paternal postpartum depression. *Journal of Affective Disorders*, *215*, 62–70. https://doi.org/10.1016/j.jad.2017.03.031

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*(3), 316–336. https://www.ncbi.nlm.nih.gov/pubmed/1918611

Coates, R., Ayers, S., & de Visser, R. (2016). Factor structure of the Edinburgh postnatal depression scale in a population-based sample. *Psychological Assessment*, *29*(8), 1016–1027. https://doi.org/10.1037/pas0000397

Cox, J. L., & Holden, J. M. (2003). *A guide to the Edinburgh postnatal depression scale*. Gaskell.
Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. *British Journal of Psychiatry*, *150*(6), 782–786. https://doi.org/10.1192/bjp.150.6.782
Dahlen, H. G., Barnett, B., Kohlhoff, J., Drum, M. E., Munoz, A. M., & Thornton, C. (2015). Obstetric and psychosocial risk factors for Australian-born and non-Australian born women and associated pregnancy and birth outcomes: A population based cohort study. *BMC Pregnancy and Childbirth*, *15*(1), 292. https://doi.org/10.1186/s12884-015-0681-2
Della Vedova, A. M., Loscalzo, Y., Giannini, M., & Matthey, S. (2022). An exploratory and confirmatory factor analysis study of the EPDS in postnatal Italian-speaking women. *Journal of Reproductive and Infant Psychology*, *40*(2), 168–180. https://doi.org/10.1080/02646838.2020.1822993
Dudeney, E., Coates, R., Ayers, S., & McCabe, R. (2023). Measures of suicidality in perinatal women: A systematic review. *Journal of Affective Disorders*, *324*, 210–231. https://doi.org/10.1016/j.jad.2022.12.091
Falissard, B. (2022). Psy: Various procedures used in psychometric (version {R package version 1.2). Retrieved from https://CRAN.R-project.org/package=psy
Fletcher, T. D. (2022). QuantPsyc: Quantitative Psychology tools (version R package version 1.6). Retrieved from https://CRAN.R-project.org/package=QuantPsyc
Giallo, R., Woolhouse, H., Gartland, D., Hiscock, H., & Brown, S. (2015). The emotional-behavioural functioning of children exposed to maternal depressive symptoms across pregnancy and early childhood: A prospective Australian pregnancy cohort study. *European Child and Adolescent Psychiatry*, *24*(10), 1233–1244. https://doi.org/10.1007/s00787-014-0672-2
Gollan, J. K., Wisniewski, S. R., Luther, J. F., Eng, H. F., Dills, J. L., Sit, D., Ciolino, J. D., & Wisner, K. L. (2017). Generating an efficient version of the Edinburgh postnatal depression scale in an urban obstetrical population. *Journal of Affective Disorders*, *208*, 615–620. https://doi.org/10.1016/j.jad.2016.10.013
Harrell, F. E., Jr. (2022). Hmisc: Harrell Miscellaneous (Version R package version 4.7-2). Retrieved from https://CRAN.R-project.org/package=Hmisc
Hollins Martin, C. J., & Martin, C. R. (2014). Development and psychometric properties of the birth satisfaction scale-Revised (BSS-R). *Midwifery*, *30*(6), 610–619. https://doi.org/10.1016/j.midw.2013.10.006
Horakova, A., Noskova, E., Svancer, P., Marcianova, V., Koliba, P., & Sebela, A. (2022). Accuracy of the Edinburgh postnatal depression scale in screening for major depressive disorder and other psychiatric disorders in women towards the end of their puerperium. *Ceska Gynekologie*, *87*(1), 19–26. https://doi.org/10.48095/cccg202219
Howard, L. M., Flach, C., Mehay, A., Sharp, D., & Tylee, A. (2011). The prevalence of suicidalideation by the Edinburgh postnatal depression scale in postpartum women in primary care: Findings from the RESPOND trial. *BMC Pregnancy and Childbirth*, *11*(1), 57. https://doi.org/10.1186/1471-2393-11-57
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
The International Consortium for Health Outcome Measurement. (2017). *Pregnancy and Childbirth Standard Set and Reference Guide, Version 1.0.3*. Retrieved from http://www.ichom.org/medical-conditions/pregnancy-and-childbirth/
Jomeen, J., & Martin, C. R. (2005). Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. *Journal of Reproductive and Infant Psychology*, *23*(2), 143–154. https://doi.org/10.1080/02646830500129297
Jomeen, J., & Martin, C. R. (2007). Replicability and stability of the multidimensional model of the Edinburgh Postnatal Depression Scale in late pregnancy. *Journal of Psychiatric and Mental Health Nursing*, *14*(3), 319–324. https://doi.org/10.1111/j.1365-2850.2007.01084.x
Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. Retrieved from https://CRAN.R-project.org/package=semTools
Kline, P. (1994). *An easy guide to factor analysis*. Routledge.

Kozinszky, Z., Toreki, A., Hompoth, E. A., Dudas, R. B., & Nemeth, G. (2017). A more rational, theory-driven approach to analysing the factor structure of the Edinburgh Postnatal Depression Scale. *Psychiatry Research*, *250*, 234–243. https://doi.org/10.1016/j.psychres.2017.01.059

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, *2*(1), 49–55.

Malouf, R., Harrison, S., Burton, H. A. L., Gale, C., Stein, A., Franck, L. S., & Alderdice, F. (2022). Prevalence of anxiety and Post-traumatic Stress (PTS) among the parents of babies admitted to neonatal units: A systematic review and meta-analysis. *EClinicalMedicine*, *43*, 101233. https://doi.org/10.1016/j.eclinm.2021.101233

Martin, C. R., Hollins Martin, C. J., Burduli, E., Barbosa-Leiker, C., Donovan Batson, C., & Fleming, S. E. (2018). The Birth Satisfaction Scale – Revised (BSS-R): Should the subscale scores or the total score be used? *Journal of Reproductive and Infant Psychology*, *36*(5), 1–6. https://doi.org/10.1080/02646838.2018.1490498

Martin, C. R., & Redshaw, M. (2018). Establishing a coherent and replicable measurement model of the Edinburgh postnatal depression scale. *Psychiatry Research*, *264*, 182–191. https://doi.org/10.1016/j.psychres.2018.03.062

Matthey, S. (2008). Using the Edinburgh postnatal depression scale to screen for anxiety disorders. *Depression and Anxiety*, *25*(11), 926–931. https://doi.org/10.1002/da.20415

Matthey, S., Fisher, J., & Rowe, H. (2013). Using the Edinburgh postnatal depression scale to screen for anxiety disorders: Conceptual and methodological considerations. *Journal of Affective Disorders*, *146*(2), 224–230. https://doi.org/10.1016/j.jad.2012.09.009

Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability Research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via monte carlo simulations. *Journal of Intelligence*, *3*(1), 2–20. https://doi.org/10.3390/jintelligence301000

Nagle, U., Naughton, S., Ayers, S., Cooley, S., Duffy, R. M., & Dikmen-Yildiz, P. (2022). A survey of perceived traumatic birth experiences in an Irish maternity sample - prevalence, risk factors and follow up. *Midwifery*, *113*, 103419. https://doi.org/10.1016/j.midw.2022.103419

Nakic Rados, S., Martinic, L., Matijas, M., Brekalo, M., & Martin, C. R. (2022). The relationship between birth satisfaction, posttraumatic stress disorder and postnatal depression symptoms in Croatian women. *Stress and Health*, *38*(3), 500–508. https://doi.org/10.1002/smi.3112

National Institute for Health and Care Excellence. (2015). *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance (Cg192)*. Retrieved from http://www.nice.org.uk/guidance/cg192

O'Hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression—a meta-analysis. *International Review of Psychiatry*, *8*(1), 37–54. https://doi.org/10.3109/09540269609037816

Phillips, J., Charles, M., Sharpe, L., & Matthey, S. (2009). Validation of the subscales of the Edinburgh postnatal depression scale in a sample of women with unsettled infants. *Journal of Affective Disorders*, *118*(1–3), 101–112. https://doi.org/10.1016/j.jad.2009.02.004

Pope, C. J., Xie, B., Sharma, V., & Campbell, M. K. (2013). A prospective study of thoughts of self-harm and suicidal ideation during the postpartum period in women with mood disorders. *Archives of Women's Mental Health*, *16*(6), 483–488. https://doi.org/10.1007/s00737-013-0370-y

Prins, A., Bovin, M. J., Kimerling, R., Kaloupek, D. G., Marx, B. P., Pless Kaiser, A., & Schnurr, P. P. (2015). *Primary Care PTSD Screen for DSM-5 (PC-PTSD-5)*. Retrieved from https://www.ptsd.va.gov

Ratislavova, K., Hendrych Lorenzova, E., Hollins Martin, C. J., & Martin, C. R. (2022). Translation and validation of the Czech Republic version of the Birth Satisfaction Scale-Revised (BSS-R). *Journal of Reproductive and Infant Psychology*, 1–17. https://doi.org/10.1080/02646838.2022.2067837

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reichenheim, M. E., Moraes, C. L., Oliveira, A. S., & Lobato, G. (2011). Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): Empirical evidence for a general factor. *BMC Medical Research Methodology*, *11*(1), 93. https://doi.org/10.1186/1471-2288-11-93

Revelle, W. (2022). Psych: Procedures for Psychological, psychometric, and personality Research (version R package version 2.2.9). Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sakalidis, V. S., Rea, A., Perrella, S. L., McEachran, J., Collis, G., Miraudo, J., Prosser, S. A., Gibson, L. Y., Silva, D., & Geddes, D. T. (2022). Longitudinal changes in wellbeing amongst breastfeeding women in Australia and New Zealand during the COVID-19 pandemic. *European Journal of Pediatrics*, *181*(10), 3753–3766. https://doi.org/10.1007/s00431-022-04580-y

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modelling* (3rd ed.). Routledge/Taylor and Francis Group.

Scottish Intercollegiate Guidelines Network. (2012). *Management of Perinatal Mood Disorders (SIGN publication no. 127)*. Edinburgh.

Smith-Nielsen, J., Egmose, I., Wendelboe, K. I., Steinmejer, P., Lange, T., & Vaever, M. S. (2021). Can the Edinburgh postnatal depression scale-3A be used to screen for anxiety? *BMC Psychology*, *9*(1), 118. https://doi.org/10.1186/s40359-021-00623-5

Steiger, J. H., & Lind, J. (1980). *Statistically-based tests for the number of common factors*. Paper presented at the Annual Spring Meeting of the Psychometric Society, Iowa City, USA.

Tuohy, A., & McVey, C. (2008). Subscales measuring symptoms of non-specific depression, anhedonia, and anxiety in the Edinburgh postnatal depression scale. *British Journal of Clinical Psychology*, *47*(Pt 2), 153–169. https://doi.org/10.1111/j.2044-8260.2008.tb00463.x

Yakupova, V., & Suarez, A. (2022). Postpartum PTSD and birth experience in Russian-speaking women. *Midwifery*, *112*, 103385. https://doi.org/10.1016/j.midw.2022.103385