Author's response to reviews

Title: Development of the Tilburg Pregnancy Distress Scale: the TPDS

Authors:

Victor J.M. Pop (V.J.M.Pop@uvt.nl)
Antoinette M. Pommer (A.M.Pommer@uvt.nl)
Monica Pop-Purceleanu (M.Pop-Purceleanu@psy.umcn.nl)
Hennie A.A. Wijnen (HAAWijnen@onsbrabantnet.nl)
Veerle Bergink (V.Bergink@erasmusmc.nl)
Frans Pouwer (F.Pouwer@uvt.nl)

Version: 2 Date: 1 September 2011

Author's response to reviews: see over
Dear Editor,

Please find enclosed a revised version of our paper: Development of the Tilburg Pregnancy Distress Scale: the TPDS

For purpose of clarity we refer in the text to the comments of the reviewers. As you will see we took over most of the comments of the reviewers and we feel that our manuscript has been markedly improved now.

We hope that our revised version is suitable for publication in your journal.

Sincerely yours,

On behalf of the authors.

Reviewer: Wai Kai Hou

Ad. 1 and 2: Throughout the text, the authors highlight the importance and advantage of using in-depth interview to generate relevant scale items for further validation. However, the authors did not provide the method of analyzing the qualitative data (e.g., content analysis, grounded theory approach) and the themes of major findings from the analysis. As a result, the audiences might not be convinced that the items generated in this study are more relevant and appropriate than those generated based on practitioners / researchers’ experience.

Second, the conceptual underpinning of the confirmatory factor analysis requires further clarification. Relative to EFA, CFA is more concept driven: researchers have a hypothesized model in mind, test the validity through EFA, and then modify and confirm the model in CFA. The authors did not explain whether they drafted the items with reference to any concepts / frameworks; therefore, the resulting CFA appeared to be an ad-hoc model that was generated based entirely on the EFA data. This mirrors the deficit of information about the major findings in the qualitative analysis: did the qualitative findings reveal that NA and partner involvement are the two major concerns among these ladies?

In response to these comments the text has been adapted as follows (lines 94-109):

The intention was to create a scale primarily based on the experiences of pregnant women since it is well known that psychological distress is often poorly recognized by clinicians (Gilboel, House and Sheldon 2005). Therefore no preliminary models were formulated and only double items were removed to construct the first draft of the questionnaire in order to lose as little information as possible. Based on the panel’s consensus, a total of 22 items were derived from an original sample of 34 candidate questionnaire items for further pilot testing. To avoid a neutral response category the items were then formatted on a four-point scale (ranging from 0= “very often” to 3 = “rarely/never”). This version was subsequently distributed in 11 community midwife offices to examine its psychometric properties (study I, sample I). These analyses were then utilized to generate a more refined version of the TPDS, which was then distributed in 10 community midwife offices who did not participate in study I. Data from study II (sample II) were finally used to conduct a confirmatory factor analysis (CFA) and to determine the concurrent and construct validity (see figure 1).
In addition, as might not have been stated clearly, the EFA and CFA were carried out in two different populations, now called sample I and sample II.

Ad. 3 The authors start off with a broad definition of the term emotional distress – stress, depression, anxiety. Then the authors go on to use an even broader term distress, to refer to the theme of the scale. However a close examination of the scale reveals that the items address emotional distress pertinent to the pregnancy or postpartum period, making up a subscale akin to the emotional distress subscale in the Revised Illness Perception Questionnaire (Moss-Morris et al., 2002). The authors might consider revising the terms that were used to refer to the theme of the scale or the factor in the EFA and CFA.

We agree with the reviewer and we took over the term he suggested in one of his comments (see below ad 9): psychological functioning in pregnant women.

Ad. 4 The authors need to provide the inclusion and exclusion criteria for recruiting the participants.

In response to these comments the text has been adapted as follows (lines 112-119):

Over a period of four months, 419 pregnant women visiting their midwife for antenatal care were invited to participate in the first pilot stage of the study, exclusion criteria were not being Caucasian, not being able to read Dutch sufficiently. Of these women, 295 (70%) consented to participate in study I (sample I). Subsequently, the second test version of the TPDS was distributed in another group of pregnant women with the same exclusion criteria used in sample I: 454 women were approached and 304 (67%) agreed to participate in study II (sample II). The women participating in each study had similar characteristics (Table 1).

Ad. 5 Combined PCA and varimax rotation are used for data reduction, testing instruments that are constructed without a prior conceptual framework or prior information (Tabachnick & Fidell, 2001, 2007). This contradicts the availability of in-depth qualitative findings in the study. The authors need to explain why the alternative methods such as factor analysis and direct oblimin rotation (assuming / allowing correlations between factors) were not used. In fact, NA and partner involvement are hardly uncorrelated among pregnant women.

In response to this comment an oblimin rotation was done with mostly the same results as with varimax rotation. Only one question that did load on the TPDS-PI factor with varimax rotation did not load when using the oblimin rotation. All analysis were done again, leaving this question out of the analysis, differences were minimal.

Ad. 6 Predictive validity refers to the extent to which the scores on an instrument are associated with the scores on a criterion measure that will be obtained at a later time. Therefore, the analysis did not address predictive validity of the TPDS. The authors need to revise the title and clarify the validity that they are testing.

‘Predictive validity’ is replaced by ‘construct validity’.
**Ad. 7** The five items that loaded on neither factors were discarded without further investigation. The authors need to discuss the content of the five items and provide reasons for the possible irrelevance or inappropriateness of the five items.

*The following sentence was added to the manuscript (lines 201-202):*

*Closer examination of these items revealed that some items were too general and others often yielded the same answers in most women.*

**Ad. 8** The authors need to provide other commonly used conventional fit indices including incremental fit index (IFI) and standardized root mean square residual (SRMR), or explain why IFI and SRMR are not used. In addition, the authors need to explain why the model is accepted though some of the reported indices are below the conventional cut-off (NFI < .90 and that for RMSEA > .05).

*We partly agree.*

*We repeated the analysis with Lisrel which provided us the SRMR showing with 0.07 that our model had a good fit.*

*We wish to disagree with the reviewer that also the IFI should be mentioned. Both prof. G. van Heck and prof. I. Komproe, experts in developing questionnaires agreed that this index (IFI) is hardly ever given in a paper describing the development of a new scale.*

**Ad. 9** The authors need to provide the findings of previous studies on the impact of partner involvement on psychological functioning in pregnant women.

*We wish to disagree with the reviewer. There are no papers describing partner involvement as an important aspect of psychological functioning in pregnant women (as we mentioned in the paper: ‘Interestingly, the current TPDS analyses indicated that perceived partner involvement (TPDS-PI sub-scale) constitutes a critically important variable for women during and after pregnancy (items 2, 4, 8, 16).’)*

**Ad. 10** The authors need to provide the theoretical and empirical literature supporting the hypothesis that TPDS is trimester specific.

*We wish not to include this remark in the changes made in the manuscript.*

*In the manuscript it is described that one may argue that the TPDS is trimester specific, however, since the Cambridge Worry Scale does not seem to be trimester specific, we do not expect the TPDS to be trimester specific either (lines 294-298).*

*‘One may argue that the scores on the TPDS are trimester specific which would in turn call for future research to validate the TPDS per trimester. However, as far as the NA-subscale is concerned, the Cambridge Worry Scale scores at different trimesters showed to be highly inter-correlated with appropriate validity at each trimester.’*

**Ad. 11** Reference(s) is needed to support the definition of the term emotional distress.

*This comment is no longer applicable since the term ‘emotional distress’ has been replaced by the term ‘psychological functioning’ as recommended in Ad 3.*

**Ad. 12** Statistical Package, not Social Package
‘Social’ has been changed in ‘Statistical’

Ad. 13 Unclear what ‘other dimensions’ means

‘Other dimensions’ is further explained (lines 251-256):

‘Although the overall TPDS and its NA sub-scale were only moderately correlated with well-recognized measures of depression (EPDS) and anxiety (GAD-7), indicating that the TPDS and its subscale ‘NA’ also assessed dimensions other than depression and anxiety, encouraging construct validity features may be derived from our finding that women with a previous diagnosis of depression/anxiety were at high risk for developing depressive and/or anxiety symptoms during pregnancy.’

Reviewer: Colin Martin

Ad. 1 One important paper that was missing in your literature review was ‘The Oxford Worries about Labour Scale (Redshaw et al. 2009). This instrument was developed in response to some of the shortcomings of the Cambridge Worry Scale, so it makes sense to mention this instrument, to ensure a comprehensive review.

We wish to agree, this reference has been added to the introduction of this paper (lines 77-78)

Ad. 2 Including as a figure the best-fit CFA model with standardized item-factor loadings would be of interest.

The figure as been added

Reviewer: Julie Jomeen

Ad. 1 The use of focus-groups as opposed to interviews to extract issues of relevance would benefit from greater justification

The use and importance of focus-groups is further explained in lines 94-97:

‘The intention was to create a scale primarily based on the experiences of pregnant women since it is well known that psychological distress is often poorly recognized by clinicians (Gilboel, House and Sheldon, 2005).’

Ad. 2 The framework through which the ‘recorded texts were evaluated should be clearly stated and explained.

The qualitative analysis is further explained in lines 94-103:

‘The intention was to create a scale primarily based on the experiences of pregnant women since it is well known that psychological distress is often poorly recognized by clinicians (Gilboel, House and Sheldon, 2005). Therefore no preliminary models were formulated and
only double items were removed to construct the first draft of the questionnaire in order to lose as little information as possible. Based on the panel’s consensus, a total of 22 items were derived from an original sample of 34 candidate questionnaire items for further pilot testing. To avoid a neutral response category the items were then formatted on a four-point scale (ranging from 0 = “very often” to 3 = “rarely/never”).

**Ad. 3** How many potential items were originally defined prior to the reduction to 22 final items?

Thirty-four items were originally defined prior to the reduction to 22 final items, this has been added in line 99:

‘Based on the panel’s consensus, a total of 22 items were derived from an original sample of 34 candidate questionnaire items for further pilot testing. To avoid a neutral response category the items were then formatted on a four-point scale (ranging from 0 = “very often” to 3 = “rarely/never”).’

**Ad. 4** What is the rationale for the choice of a 4 point scale?

The rationale for choosing a 4 point scale has been added in lines 102-103.

‘Based on the panel’s consensus, a total of 22 items were derived from an original sample of 34 candidate questionnaire items for further pilot testing. To avoid a neutral response category the items were then formatted on a four-point scale (ranging from 0 = “very often” to 3 = “rarely/never”).’

**Ad. 5** Should the first line read ‘… visiting their midwife for antenatal care’, not ‘… antenatal control’

The line has been changed.