Out of the box : moving from categories to dimensions in the phenomenology of depression and anxiety.

Hollander-Gijsman, M.E. den

Author: Hollander-Gijsman, Margien Elisabeth den
Title: Out of the box : moving from categories to dimensions in the phenomenology of depression and anxiety
Issue Date: 2013-12-11

CHAPTER 6

DISTINGUISHING SYMPTOM DIMENSIONS OF DEPRESSION AND ANXIETY: AN INTEGRATIVE APPROACH

Journal of Affective Disorders 2012; 136, 693-701

Margien E. den Hollander-Gijsman
Klaas J. Wardenaar
Edwin de Beurs
Nic J.A. van der Wee
Ab Mooijaart
Stef van Buuren
Frans G. Zitman

Abstract

Background

Clark and Watson developed the tripartite model in which a symptom dimension of 'negative affect' covers common psychological distress that is typically seen in anxious and depressed patients. The 'positive affect' and 'somatic arousal' dimensions cover more specific symptoms. Although the model has met much support, it does not cover all relevant anxiety symptoms and its negative affect dimension is rather unspecific. Therefore, we aimed to extend the tripartite model in order to describe more specific symptom patterns with unidimensional measurement scales.

Method

1333 outpatients provided self report data. To develop an extended factor model, exploratory factor analysis (EFA) was conducted in one part of the data (n=578). Confirmatory factor analysis (CFA) was conducted in the second part (n=755), to assess model-fit and comparison with other models. Rasch analyses were done to investigate the unidimensionality of the factors.

Results

EFA resulted in a 6-factor model: feelings of worthlessness, fatigue, somatic arousal, anxious apprehension, phobic fear and tension. CFA in the second sample showed that a 6-factor model with a hierarchical common severity factor fits the data better than alternative 1- and 3-factor models. Rasch analyses showed that each of the factors and the total of factors can be regarded as unidimensional measurement scales.

Limitations

The model is based on a restricted symptom-pool: more dimensions are likely to exist.

Conclusion

The extended tripartite model describes the clinical state of patients more specifically. This is relevant for both clinical practice and research.

6.1 Introduction

The traditional distinction between depressive and anxiety disorders has often been challenged for several reasons. First, high rates of comorbidity between depression and anxiety disorders are suspected to be an artifact of this distinction (Brown et al., 2001; De Graaf et al., 2002; Kessler et al., 1996). Second, depression and anxiety have overlapping key-symptoms, rendering depression- and anxiety measures highly correlated and only modestly discriminative (Clark & Watson, 1991). Third, the diagnoses encompass heterogeneous disorders. For instance, two patients with a similar diagnosis of Major Depressive Disorder (MDD) only have to share one out of nine criterion-symptoms, making the label MDD very unspecific (Widiger & Samuel, 2005). As a consequence, specific etiological effects are hard to detect in research because of the large variability (noise) within diagnosis groups. Fourth, the use of dichotomous criteria with arbitrary boundaries leaves us with a many subsyndromal subjects, whose etiology and risk profile are often highly similar to patients with full-fledged disorders (De Beurs et al., 1999). Fifth, using dichotomous diagnoses in research reduces statistical power, increasing the need for larger sample sizes (MacCallum, Zhang, Preacher, & Rucker, 2002).

One often proposed way to overcome these problems is the use of a dimensional approach (Clark, 2005; Cuthbert, 2005; Krueger et al., 2005). Dimensions represent continua of increasing severity on different symptom-domains (Goldberg, 2000) and an individual's clinical state is described with a pattern of specific dimensional scores. Dimensions circumvent comorbidity, describe a patients' clinical state specifically and cover the full spectrum of severity from healthy to pathological.

Several dimensional approaches to depression and anxiety have been proposed. Well known is the *tripartite model* (Clark and Watson, 1991), which consists of 3 dimensions. The 'negative affect' (NA) dimension covers general psychological distress symptoms, common to both depressive- and anxiety disorders and could account for their observed overlap and comorbidity. The 'positive affect (PA)' dimension covers the symptoms of anhedonia (e.g. lack of enthusiasm and excitement), specific for depression. The 'somatic arousal (SA)' dimension covers symptoms of somatic hyperarousal, specific for anxiety. Although the tripartite model has been found to be structurally valid, SA has been shown to be mainly specific for panic disorder (Mineka et al., 1998). Hence, several model extensions have been proposed to better account for the heterogeneity of anxiety (Chorpita, 2002; Joiner & Lonigan, 2000; Mineka et al., 1998).

Another model devised to do more justice to the internal heterogeneity of anxiety is the *valence-arousal model* (Heller, Nitschke, Etienne, & Miller, 1997). In this model, a distinction is made between two underlying anxiety factors: 'anxious apprehension' and 'anxious arousal', the latter resembling the SA dimension of the tripartite model. Anxious apprehension is an additional factor that is characterized by a concern for the future and verbal rumination about negative expectations and fears and is hypothesized to play an important role in the etiology of anxiety (Nitschke, Heller, Palmieri, & Miller, 1999).

Although both models have contributed to the field considerably, neither model was fully supported across different lines of research. Concluding a review on the various models for depressive and anxiety disorders, Shankman and Klein (2003) stated that a model with two to four dimensions might not be sufficient to do justice to all relevant common and discrete symptoms of anxiety and depression. However, the validity and usability of specific aspects of both the tripartite model and the valence-arousal model were supported.

Mineka and colleagues (1998) proposed a hierarchical model in which psychopathology was defined by a common, overarching factor of negative affect and specific lower-order factors describing the unique components of mood- and anxiety disorders. They proposed that SA could be seen as specific to panic disorder and that additional dimensions could account for distinct symptoms of other anxiety disorders. They suggested to "view individual disorders as representing unique combinations of different types of symptoms, with each type showing varying degrees of non-specificity and with no type being entirely unique to any single disorder" (Mineka et al., 1998). Several studies referring to this hierarchical model used the DSM-IV diagnoses as unit of research. They assumed that all lower level dimensions corresponded to different DSM-IV diagnoses (Krueger, 1999; Vollebergh et al., 2001; Watson, 2005). These studies presented hierarchical models based on DSM-categories and were effective in presenting a partial explanation of the high rates of comorbidity between depression and anxiety in the DSM-IV. Another way to operationalize the hierarchical model is by developing a model with dimensions for unique symptoms of specific mood- and anxiety disorders in addition to common symptom scales (Mineka et al., 1998). In previous work we presented a proposal for such an extension of the tripartite model in which each of five dimensions was more or less specific for one or more disorders (Den Hollander-Gijsman et al., 2010). Several studies have shown such an approach to work well (Simms, Gros, Watson, & O'Hara, 2008; Simms, Prisciandaro, Krueger, & Goldberg, 2012).

Due to the above-described problems with the DSM, it is likely that dimensions do not follow the strict divisions of the DSM-IV. Therefore, a dimensional model should primarily describe the unique profiles of individuals rather than of DSM-disorders. Consequently, dimensions should therefore be based on more objective criteria such as one-dimensionality, discriminative ability (between individuals) and external validation, e.g., with biological markers. Almost all abovementioned work was conducted with factor-analyses. It is often overlooked that these analyses only inform about underlying structures of data and do not imply that individual factors are unidimensional. To determine the latter, additional Rasch analyses should be conducted to check if and how the items are lined up along an underlying severity dimension (Wright & Masters, 1982). Only if a factor fits to the Rasch model, it can be regarded as a dimension with a valid additive measurement scale. This is essential if we wish to define psychopathology with dimensions.

The current study was aimed to integrate aspects of the abovementioned models into one broad dimensional model, without taking DSM-IV diagnoses as a point of departure or specificity to particular DSM-IV diagnoses as a sign of validity. Instead, we aimed for a multidimensional model to characterize individual patients in terms of their specific symptom profile. As point of departure we used a large item-pool that included (1.) the items of the Mood and Anxiety Symptom Questionnaire (Watson & Clark, 1991) to measure NA, PA and SA, (2.) items of the Brief Symptom Inventory (Derogatis, 1975), to measure fearfulness and (3.) newly designed items to measure anxious apprehension. Several analyses were conducted in two large samples (n=578 and n=755) of psychiatric outpatients. The underlying factor-structure of the item pool was explored using exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) in the first sample. In the second sample CFA was used to evaluate the fit of this structure and compare it with alternative models: a one factor model, a three factor model (the tripartite model), a higher-order model and a bifactor hierarchical model. Finally, Rasch analyses were performed to investigate and improve the unidimensionality of each factor and to evaluate whether they could be used as reliable additive subscales.

6.2 Methods

6.2.1 Participants and Procedure

This study was conducted on data collected through Routine Outcome Monitoring (De Beurs et al., 2011). ROM is a monitoring system for patient care, implemented in the outpatient clinics of Rivierduinen Psychiatric Hospital (a large organization for the provision of mental health care in the province of Zuid-Holland, the Netherlands) and the psychiatric department of the Leiden University Medical Center (LUMC). All outpatients referred to these clinics by their general practitioner for treatment of a mood-, anxiety- or somatoform disorder have an assessment session with a psychiatric research nurse at the start of treatment. During this session a standardized diagnostic interview, rating scales, and self-report rating instruments are administered. Two patient samples were composed of respectively 578 and 755 outpatients, who had paid their first visit to the clinic between March 2005 and June 2006 and had been assessed with Routine Outcome Monitoring.

6.2.2 Measures

6.2.2.1 Mini International Neuropsychiatric Interview (M.I.N.I.) Plus 5.0.0.-R

The M.I.N.I. is a short structured diagnostic interview developed to explore the presence of 23 Axis-I disorders according to the DSM-IV diagnostic criteria (Sheehan et al., 1998). In this study the Dutch translation of the M.I.N.I.-Plus 5.0.0-R (Van Vliet et al., 2000) was used to screen for the presence of current disorders. Psychiatric research nurses who were extensively trained and supervised performed the interviews.

6.2.2.2 Mood and Anxiety Symptom Questionnaire (MASQ)

The MASQ was used to assess the severity of symptoms of depression and anxiety over the past week (De Beurs et al., 2007 (Dutch version); Watson & Clark, 1991). The MASQ consists of 90 items, divided into 5 subscales measuring different aspects of the tripartite model: 1) anhedonic depression; 2) anxious arousal; 3) general distress depression; 4) general distress anxiety, and 5) general distress mixed. All items are rated on a 5-point rating scale (1 [not at all] to 5 [very much]). All items of the MASQ denoting positive feelings (anhedonic depression scale) were reversed keyed before analysis to make the interpretation of the results more straightforward.

6.2.2.3 Brief Symptom Inventory (BSI)

The Brief Symptom Inventory (De Beurs, 2005; Derogatis & Melisaratos, 1983) is a shortened version of the Symptom Checklist (SCL-90) (Derogatis et al., 1973), and was used to measure psychological complaints or symptoms. The BSI consists of 53 items that are rated on a 5-point scale (0 [not at all] to 4 [very much]). The items measure nine subscales: somatic complaints, cognitive problems, interpersonal sensitivity, depression, anxiety, hostility, phobic fear, paranoid thinking, and psychoticism.

6.2.2.4 Anxious apprehension

We formulated four self report items (AA-01 to AA-04) to measure anxious apprehension (e.g., “I worried about bad things that might happen”). To determine the face validity of these items, they were judged by two individual clinical experts (psychiatrist and psychologist).

6.2.2.5 Final item-pool

A selection was made from the BSI and the MASQ items to prevent redundancy. The items of all five MASQ subscales were included (77 out of the 90 items: the remaining items were not assigned to any subscale (Watson & Clark, 1991). From the BSI, the items of the *anxiety* and *phobic fear* subscales were selected. Together with the four items measuring anxious apprehension, this resulted in an item-pool of 91 unique items.

6.2.3 Statistical analyses

6.2.3.1 Model selection

Before the analyses, all items of the BSI were recoded from 0-4 to 1-5 to match with the scoring of the MASQ. EFA was used in sample 1 to investigate how many and which factors should be retained to model the underlying structure of the item-pool. Oblique factor rotation (oblimin) was used, because it does not assume that factors are uncorrelated. Factor extraction was done by use of a scree-plot. Items were retained for each factor if they had a high (>0.40) factor-loading and did not have a high (>0.40) loading on any of the other factors. The cut-off of 0.40 was chosen to balance between over- and under inclusion of items within each factor. The EFA was conducted using SPSS 17. Next, CFA were run to evaluate the fit of a 1-factor model on each extracted factor. Model-fit was evaluated with fit-indices (see below for the used methods and cut-off criteria). If fit was inadequate, the scale was further examined with EFA and items with low factor scores were deleted from the scale to improve fit. These steps were repeated until each factor fit well to the data.

6.2.3.2 Model evaluation: Confirmatory Factor Analyses (CFA)

To investigate the validity of the model structure that was identified in sample 1, CFA was conducted in sample 2. The newly identified multi-factor model was compared to four alternative models. In a *1-factor* model, all items loaded on one common factor. In the tripartite model the negative affect-, positive affect- and somatic arousal-related items loaded on three different factors (Clark & Watson, 1991). In a *higher order model,* a higher order severity factor loaded on all identified (Van Kampen D., 2006) factors. In a hierarchical bifactor model different sets of items loaded on specific factors and, at the same time, all items loaded on one general severity factor (following Mineka et al., 1998). In each tested model the factor-loadings were set to be freely estimated; per factor one factor-loading was fixed to one. In the result section schematic illustrations of the five models are provided.

The data were all categorical and non-normally distributed, thus maximum likelihood (ML) estimation of model-fit would likely result in underestimations of model-fit (Byrne, 2006). Therefore, we used an approach for categorical data (Bentler, 2006). First, a matrix of polychoric correlations between the items was generated. Second, model fit-statistics were estimated with ML. Third, the fit-statistics were corrected with an appropriate weight-matrix to obtain *robust* fit-statistics (Satorra & Bentler, 1988). These robust statistics have been shown to perform well for categorical and non-normal data (Byrne, 2006). The following fit-indices were used to assess model-fit: the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA) and the Akaike Information Criterion (AIC). A CFI ≥ 0.90 and a RMSEA ≤ 0.08 indicates adequate fit (Hu & Bentler, 1999). The AIC can be used to compare different models, balancing statistical goodness-of-fit and the number of model parameters; the model with the lowest AIC can be regarded as potentially most useful (Bentler, 2006). The EQS statistical package (Multivariate Software Inc., Encino, California, USA) was used to conduct the analyses.

Spearman correlations between the raw sum scores of the different factors were computed to evaluate their interrelatedness with SPSS 17.

6.2.3.3 Model evaluation: Rasch analyses

To investigate the unidimensionality of the identified factors, fit to the Rasch model was investigated in sample 2. Calculations were done with RUMM2020 (RUMM Laboratory, Perth, WA, Australia). The Rasch model assumes that the probability of a person's response on an item is described by a *logistic* function of the distance between the location of the person and the location of the item on the underlying linear *severity dimension*. If a person is located higher on the underlying dimension than an item, the probability that the person responds with the highest response option on a Likert-item is very high. On the other hand, if the person is located lower on the dimension than the item, the probability of the lowest response option is high. If a group of items fits well to the Rasch model, in theory all of the items are lined up along one underlying dimension in order of increasing severity. An important implication of adequate fit to the Rasch model is that this indicates that the ordinal responses on the items can be added up to a linear interval-scale that is a *sufficient statistic* for the underlying severity dimension, which means that the factor is a unidimensional measurement scale (Wright & Masters, 1982). The latter was why we chose to use the 1-parameter Rasch model instead of a more-parameter item response model, which allows for more subtle fit assessment but does not have a simple sufficient statistic.

The unrestricted partial credit model was used for fit-estimation. To estimate the fit to the model, the unweighted mean square standardized residual (outfit) was calculated for each item (formulas from: Wright & Masters, 1982, p100). Outfit was used because it is much less affected by large sample size because it is basically a χ^2 statistic divided by its degrees of freedom. An outfit for an item that is close to 1 and within the range of 0.7 to 1.3 is considered to indicate adequate fit (Wright & Stone, 1979). In the current analyses, the standardized residuals were calculated and outputted by RUMM and the mean residual across all persons (the outfit) was calculated for each item using Microsoft Excel. Persons with a total scale score of 0 or with fit-residuals>|2.5| were automatically excluded from all calculations because they do not behave in line with the Rasch model expectations.

For each factor, the same analytic procedure was followed to assess fit of items to the Rasch model. First, for each item the polytomous *category probability plot* was screened for disordered thresholds between response categories. If along the underlying dimension, a category always had a lower probability of endorsement than a neighboring category, the lower- probability category was 'collapsed' with the higher-probability category. If the category with a higher probability was one step down on the response scale, the lower-probability category was *collapsed down* and if the category with a higher probability was one step up on the Likert scale, the lower-probability category was *collapsed up*. Second, the fit of the items within each factor was assessed to see if fit had improved with rescoring and extra rescoring was undertaken if necessary. Third, if items fit well, differential item functioning (DIF) was used to investigate whether item-functioning differed across gender and age-tertiles. This method uses an ANOVA, which was likely to pick up less relevant DIF due to our large sample-size. Therefore, if significant DIF was found for an item, the item-locations were additionally compared across subgroups (e.g. men vs. women) to judge whether DIF was relevant and could potentially harm generalizability. Fourth, the person-separation index was calculated and the number of severity strata that could be discriminated was derived from the separation-ratio (G).

6.3 Results

6.3.1 Demographic and diagnostic characteristics

The two samples contained respectively 66 % and 61 % females and the mean age was 37 years (range 18-78) for both samples. No significant differences were found between the two samples on any of the listed demographic and psychopathology characteristics (see Table 6.1).

6.3.2 Model selection: EFA and CFA

EFA with Oblimin rotation in sample 1 yielded various feasible solutions. Based on the number of unique loading items per factor and the interpretability of the factors, we decided on a seven-factor solution with 56% of explained variance. The factors were: feelings of worthlessness, positive affect, fatigue, somatic arousal, anxious apprehension, phobic fear, and tension (eigenvalues: 31.3, 6.5, 3.7, 2.7, 2.3, 2.2, and 1.9). When qualitatively comparing this model with the tripartite model, the dimensions positive affect and anxious arousal are retained, a new dimension fatigue emerges, and the dimension negative affect is subdivided into four dimensions: feelings of worthlessness, phobic fear, anxious apprehension and tension (see Table 6.2).

The positive affect factor was entirely composed of positively formulated feelings or emotions (reverse keyed items), which suggests that these items mainly load on the same factor because of their shared response-format: a method effect rather than a truly separate concept (Russell & Carroll, 1999; Spector, Van Katwyk, Brannick, & Chen, 1997). We decided to omit this factor from further analyses to decrease the chance on bias in the model by response

Variables	Sample 1 *March 2005 – September 2005*	Sample 2 *October 2005 - June 2006*
N	578	755
Mean Age (SD)	37 (.13)	37 (.12)
Age range	(18-78)	(18-71)
Number of Females (%)	382 (66%)	463 (61%)
Mean BSI total (SD)	1.1 (0.71)	1.1 (0.69)
BSI-total range	(0-3.3)	(0-3.6)
Diagnoses (%):		
Depression/dysthymia	266 (46%)	333 (44%)
Anxiety disorder	273 (47%)	352 (47%)
Somatoform disorder	101 (18%)	102 (14%)
Diagnostic groups (%)		
No depression and anxiety	166 (29%)	210 (28%)
Only anxiety disorder	146 (25%)	212 (28%)
Only depressive disorder	139 (24%)	193 (26%)
Depressive and anxiety disorder	127 (22%)	140 (19%)

BSI = Brief Symptom Inventory

Table 6.1 Demographic and psychopathology characteristics for sample 1 (n=578) and sample 2 (n=755).

format. Due to this decision, all dimensions in our model are measured with negatively formulated items only. To do justice to the construct ‘positive affect/positive activation’, we preserve the factor ‘fatigue’ because in theory, positive affect and fatigue can be interpreted as opposite poles of the same dimension (Clark & Watson, 1991, p. 321).

For each remaining factor, all items with a substantial (>0.40) factor-loading were retained (feelings of worthlessness: 5 items, fatigue: 8 items, somatic arousal: 13, anxious apprehension: 5 items, phobic fear: 4 items, and tension: 6 items). CFA with each of these factors showed that a one-factor model fit the factors ‘feelings of worthlessness’ (CFI=0.98), Tension (CFI=0.98), and ‘phobic fear’ (CFI=0.97) very well. For the other three factors, model-fit was inadequate (CFI ranged from 0.80 to 0.86). Therefore, an additional EFA was done on each of these three factors to select the items with the highest loadings on the factor. Subsequent CFA’s showed these fine-tuned factors to have satisfactory fit to a

Tripartite model	New dimensions
Somatic arousal	Somatic arousal
Positive affect	Fatigue
	Positive affect*
Negative affect	Phobic fear
	Anxious apprehension
	Feelings of worthlessness
	Tension

* Positive affect was entirely composed of positively formulated items and we decided to continue the analyses without this dimension.

Table 6.2 Comparison of the new factors with the dimensions of the tripartite model.

one-factor model (CFI ranged from 0.95 to 1.00). For anxious apprehension (AA) and phobic fear (PF) the RMSEA was greater than 0.1 (.211 and .137 respectively) and we did not succeed to reduce these values with further modifications to the factors.

6.3.3 Model evaluation: CFA

Schematic illustrations of the five models are depicted in Figure 6.1, and the results of the CFA in sample 2 are shown in Table 6.4. The newly identified 6-factor model (model 3) showed adequate fit (CFI=0.95; RMSEA=0.081). The 1 factor model (model 1) resulted in worse model-fit (CFI=0.89; RMSEA=0.13). To test fit to the tripartite model, the items representing feelings of worthlessness, tension, anxious apprehension, and phobic fear were taken together in one NA factor. Together with the fatigue factor and the SA factor, these formed the 3-factor tripartite model (model 2). This model fit worse than the 6-factor model (CFI=0.92; RMSEA=0.11). A 6-factor model, with a higher-order factor (model 4) fits better than the regular 6-factor model (CFI=0.99; RMSEA=0.046) and the bifactor hierarchical 6-factor model (model 5) showed the best fit (CFI=0.99; RMSEA=0.043). In addition this model had the lowest AIC (130.8) compared to the other models (AIC range: 200.3 to 4177.9). This indicated that the best model to describe the underlying structure of our data-pool has 6 different factors with one additional overarching severity factor.

6.3.4 Intercorrelations

The correlations between the sum scores of each of the six factors in sample

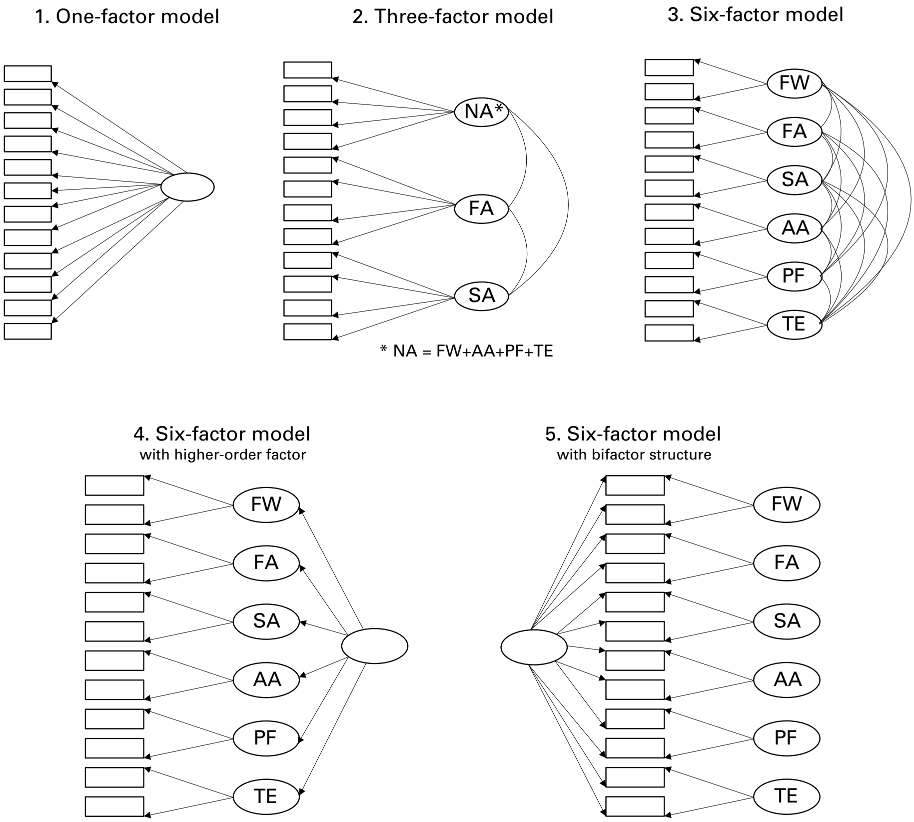

Figure 6.1 Schematic illustration of a priori structural models. The 5 pictures depict only a few relevant parameters; the exact number of items and error terms on symptoms and intermediate factors are omitted for clarity.
NA=negative affect, FW=feelings of worthlessness, FA=fatigue, SA=somatic arousal, AA=anxious apprehension, PF=phobic fear, TE=tension

2 are displayed in Table 6.3. The coefficients ranged from 0.30 to 0.63. The correlations between fatigue and phobic fear (r=0.30), between feelings of worthlessness and somatic arousal (r=0.33) and between somatic arousal and phobic fear (r=0.33) were all modest. The correlations between fatigue and tension (r=0.63), between feelings of worthlessness and anxious apprehension (r=0.59) and between tension and anxious apprehension (r=0.58) were high. All other correlations ranged from 0.37 to 0.56. This indicates that the identified structure consists of moderately to strongly related constructs.

	FW	FA	SA	AA	PF	TE
Feelings of worthlessness (FW)	1.00	-	-	-	-	-
Fatigue (FA)	0.54	1.00	-	-	-	-
Somatic Arousal (SA)	0.33	0.47	1.00	-	-	-
Anxious Apprehension (AA)	0.59	0.43	0.43	1.00	-	-
Phobic Fear (PF)	0.42	0.30	0.33	0.39	1.00	-
Tension (TE)	0.56	0.63	0.55	0.58	0.37	1.00

All correlation coefficients significant at $p<0.01$

Table 6.3 Spearman correlations for the new scales (sumscores) in sample 2 (n=755).

6.3.5 Model evaluation: Rasch analyses

Rasch analyses (Table 6.5) were performed for the complete item-set and for the different factors that were identified using EFA and CFA.

6.3.5.1 All items

Because we found an overarching general severity factor we investigated the fit of the Rasch model on all items within the identified model. Because most items appeared to have disordered thresholds, they were recoded to a 4-point scale (0,1,1,2,3). Items BSI08 and BSI28 were recoded to (0,0,0,1,1), BSI31 and BSI43 to (0,0,1,1,2) and MASQ79 and MASQ81 were recoded to (0,1,1,1,2). Outfit ranged from 0.73 to 1.42 and only two items had an outfit that exceeded the criteria for good fit (BSI31: outfit=1.42; and BSI43: outfit=1.31). The person-separation index was 0.93, which indicated that the scale could be used to discriminate between five severity strata ($G\approx4$, Wright and Masters, (1982)).

6.3.5.2 Feelings of Worthlessness

In the feelings of worthlessness factor, adequate threshold ordering was obtained by rescoring all items to a 3-point scale (0,0,1,1,2). Outfit ranged from 0.72 to 0.92, indicating adequate fit to the Rasch model. No DIF was found. The person-separation index was 0.84, which indicated that the scale can be used to discriminate between 3 severity strata ($G\approx2$).

6.3.5.3 Fatigue

In the fatigue factor, adequate threshold ordering was obtained by rescoring all items to a 4-point scale (0,1,1,2,3). Outfit ranged from 0.72 to 0.89, indicating adequate fit to the Rasch model. No DIF was found. The person-separation index was 0.84, which indicated that the scale can be used to discriminate between 3 severity strata (G≈2).

6.3.5.4 Somatic Arousal

In the somatic arousal factor, adequate threshold ordering was obtained by rescoring all items to a 3-point scale: MASQ48 and MASQ75 to (0,0,1,1,2) and MASQ79 and MASQ81 to (0,1,1,1,2). Outfit ranged from 0.70 to 0.85, indicating adequate fit to the Rasch model. DIF was found across age on item MASQ81 ('Muscles were tense or sore'): item locations ranged from -0.97 (middle age) to -0.32 (low age), indicating that item-functioning differs slightly across age groups. The person-separation index was 0.65, which indicated that the measurement scale can be used to discriminate between roughly 2 severity strata (G≈1.5).

6.3.5.5 Anxious Apprehension

In the anxious apprehension factor, adequate threshold ordering was obtained by rescoring all items to a 3-point scale: items AA-01 and AA-02 to (0,0,1,1,2) and AA-03 and AA-04 to (0,0,0,1,2). Outfit ranged from 0.65 to 1.09, indicating adequate fit to the Rasch model for only two items (AA-01 and AA-02). Two other items consistently failed to adequately fit to the model (AA-03 and AA-04), even after further rescoring. The latter items were thus dropped from the scale. No DIF was found. The remaining two items only had a person-separation index of 0.54, which indicated that the measurement scale can not be used to discriminate different strata of severity (G≈1). The factor is thus not very useful as a measurement scale.

Model	DF	S-B $\chi 2$	AIC
1. One factor	351	4879.4	4177.9
2. Three factors	350	3572.4	2872.4
3. Six factors	341	1244.4	562.4
4. Six factors (higher order)	344	1069.4	200.34
5. Six factors (bifactor)	322	774.8	130.82

Analyses based on polychoric correlation matrix; model-fit estimation with ML, Chi-square and fit indices adjusted for non-normality with Satorra-Bentler correction. S-B $\chi 2$ = Satorra-Bentler Chi-square; AIC = Akaike Information Criterion;

Table 6.4 Results of confirmatory factor analyses in sample 2 (n=755).

6.3.5.6 Phobic Fear

In the phobic fear factor, adequate threshold ordering was obtained by rescoring all items to a 2-point scale: items BSI08 and BSI28 to (0,1,1,1,1) and BSI31 and BSI43 to (0,0,1,1,1). Outfit ranged from 0.83 to 1.07, indicating adequate fit to the Rasch model. No DIF was found. The person separation index was 0.61, which indicated that the measurement scale could be used to discriminate between 2 severity strata (G≈1.5).

6.3.5.7 Tension

In the tension factor, adequate threshold ordering was obtained by rescoring all items to a 4-point scale: MASQ15, MASQ17 and MASQ77 to (0,1,1,2,3) and MASQ50, MASQ59 and MASQ82 to (0,1,2,2,3). Outfit ranged from 0.67 to 0.92, indicating adequate fit to the Rasch model for all but one item. Item MASQ79 failed to fit the Rasch model (Outfit: 0.67), even after further rescoring and was therefore dropped from the scale. DIF was found across gender on item MASQ50 ('feeling restless'): item location was slightly higher (0.46) in females than in males (0.19). However the location-difference was small (<0.50), indicating only limited influence on the generalizability of measurement. The five remaining items had a separation index of 0.80, which indicated that the measurement scale can be used to discriminate between 3 severity strata (G≈2).

6.4 Discussion

The aim of the current study was to develop a dimensional model for depression and anxiety of clearly distinguishable and easily assessable dimensions, integrating the approaches of the tripartite model, the valence-arousal model and the hierarchical model.

CFI	NFI	RMSEA	90% CI (RMSEA)
0.89	0.88	0.131	0.128 -0.134
0.92	0.91	0.111	0.107 -0.114
0.98	0.98	0.059	0.056-0.063
0.99	0.98	0.046	0.042-0.050
0.99	0.99	0.043	0.039-0.047

CFI = Comparative Fit Index; NFI = Normed Fit Index; RMSEA=Root Mean Square Error of Approximation; 90% CI (RMSEA) = 90% Confidence Interval of RMSEA.

Scale	Item number	Item content
Feelings of worthlessness	MASQ74	Was disappointed in myself
	MASQ13	Felt worthless
	MASQ24	Blamed myself for a lot of things
	MASQ64	Felt inferior to others
	MASQ47	Felt like a failure
Fatigue	MASQ90	Got tired or fatigued easily
	MASQ39	Felt like it took extra effort to get started
	MASQ56	Felt sluggish or tired
	MASQ19	Felt faint
	MASQ66	Felt really slowed down
Somatic Arousal	MASQ79	Was trembling or shaking
	MASQ81	Muscles were tense or sore
	MASQ48	Had hot or cold spells
	MASQ75	Heart was racing of pounding
Anxious apprehension	AA-03	I worried about bad things that could happen
	AA-04	I was concerned about things that could happen
	AA-01	I thought that things would end up badly for me
	AA-02	I had the feeling that something bad was going to happen
Phobic Fear	BSI43	Feeling uneasy in crowds, such as shopping or at a movie
	BSI28	Feeling afraid to travel on buses, subways, or trains
	BSI31	Having to avoid certain things, places, or activities
	BSI08	Feeling afraid in open spaces or on the streets
Tension	MASQ77	Felt tense or “high-strung”
	MASQ59	Was unable to relax
	MASQ15	Felt nervous
	MASQ17	Felt irritable
	MASQ50	Felt very restless
	MASQ82	Felt keyed up, “on edge”

BSIxx = Brief Symptom inventory items; MASQxx = mood and anxiety symptom questionnaire items; AAxx = customly developed anxious apprehension items.
Items ordered according to their location within their scale; adequate outfit coefficients printed in bold font.

Table 6.5 Results of Rasch analyses in sample 2 (n=755).

Threshold			Item location	Outfit
1	2	3		
-1.80	0.94	-	-0.43	**0.79**
-1.50	1.28	-	-0.11	**0.76**
-1.20	1.49	-	0.14	**0.92**
-1.40	1.78	-	0.19	**0.80**
-0.98	1.39	-	0.21	**0.72**
-3.49	1.52	-	-0.98	**0.79**
-3.15	2.04	-	-0.55	**0.82**
-2.67	2.72	-	0.03	**0.72**
-2.18	2.93	-	0.37	**0.89**
-1.33	3.60	-	1.13	**0.76**
-1.91	1.82	-	-0.04	**0.79**
-2.32	0.71	-	-0.80	**0.85**
-0.45	0.99	-	0.27	**0.70**
-0.14	1.30	-	0.58	**0.72**
-1.30	0.45	-	-0.42	0.65
-1.71	0.04	-	-0.83	0.65
-0.89	1.64	-	0.37	**1.09**
-0.43	2.20	-	0.88	**0.80**
-0.64	-	-	-0.64	**1.07**
0.13	-	-	0.13	**0.87**
0.19	-	-	0.19	**0.88**
0.32	-	-	0.32	**0.83**
-3.69	1.57	-	-1.06	0.67
-2.15	0.30	1.23	-0.21	**0.79**
-2.20	0.56	1.96	0.11	**0.88**
-2.16	0.51	2.01	0.12	**0.92**
-1.57	0.49	2.26	0.39	**0.76**
-0.62	0.91	1.64	0.64	**0.79**

Model-development and evaluation were performed in independent samples. In the first sample a six-factor model was identified, comprising the following factors: feelings of worthlessness, fatigue, somatic arousal, anxious apprehension, phobic fear and tension. In the second sample, confirmatory factor analyses showed that a bifactor hierarchical model with a general severity factor and six specific factors fit best to the data, compared to other models. Additional Rasch analyses showed that five of the six factors were truly one-dimensional and could be used as measurement scales. Only the anxious apprehension factor was found not to be unidimensional, although this does not imply that the identified structure is invalid. Importantly, we also found good fit of the Rasch model for all items together, which is in line with the identified bifactor structure of six specific factors and a general severity factor.

These results have some interesting implications. First, they show that a hierarchical 6-factor model is optimal to describe the structure of the symptom dimensions of mood- and anxiety disorders, when integrating important aspects of the tripartite model and the valence arousal model. As suggested by Mineka et al.(1998), the six lower order factors describe a patients' specific symptom-profile, while at the same time the complete set of items reflects overall severity. Importantly, our findings are in line with earlier studies (Simms et al., 2008; Simms et al., 2012) and lend further support to the idea that symptomatology of depression and anxiety has a hierarchical structure.
In the current model the dimension *tension* was most generic and was correlated relatively strongly with all the other dimensions in the model and could be regarded as a small, more homogeneous subfactor of NA.

Somatic arousal, anxious apprehension and phobic fear all fall into the anxiety realm. The present model thus distinguishes three distinct dimensions of symptomatology, relevant to anxiety. Both phobic fear and anxious apprehension are valuable additions to the single dimension of SA in the tripartite model, because they reflect the behavioral and the cognitive components of fear and anxiety. Both dimensions were only modestly intercorrelated ($r=.39$), indicating that they measure two distinct constructs. Phobic fear is a relevant construct because it is a defining aspect of panic disorder with agoraphobia, social phobia and specific phobia (Den Hollander-Gijsman et al., 2010). Anxious apprehension was previously found to play an important role in anxiety, as shown by imaging studies on the valence-arousal model (Heller et al., 1997). Thus, by integrating these different anxiety-related constructs the current model better accounts for the heterogeneity of anxiety.

Feelings of worthlessness and fatigue are dimensions that reflect aspects of a depressed state. According to the tripartite model, the factor fatigue which reflects loss of energy can be interpreted as the negative pole of the dimension 'positive affect/anhedonia' (Clark & Watson, 1991). Besides 'fatigue', a positive affect factor emerged in the factor analysis, including all positively formulated items in the item-pool and was thus likely to reflect a method effect. We therefore decided not to include this dimension in the analyses to evaluate the model. For future research, it would be interesting to measure both NA and PA with both positively and negatively phrased items (and in both a clinical sample and a sample from the general population).

The current study had several strong characteristics. First, the sample was large, which increases reliability, and included a broad range of outpatients with mood-, anxiety-, and somatoform disorders, assuring the generalizability of the results to the target population. Second, model-development and confirmation were conducted in independent samples, supporting replicability of the identified model. Third, in addition to investigating the factor structures, the usefulness of the factors as one-dimensional measurement scales was also evaluated with Rasch analyses (Wright & Masters, 1982).

The results should also be interpreted in the light of some limitations. First, the results only apply to outpatients with a limited range of severity and specific demographic characteristics and can thus not be directly generalized to healthy controls or inpatients. Second, model-development was based on a limited symptom-pool, which may have restricted the number of factors that was identified. In reality, even more dimensions are expected to exist, such as externalizing dimensions (Krueger, Markon, Patrick, & Iacono, 2005) comprising concepts such as ‘anger’ or ‘aggression’ (Pasquini, Picardi, Biondi, Gaetano, & Morosini, 2004; Picardi, Morosini, Gaetano, Pasquini, & Biondi, 2004). Third, although the current study is based on a strong combination of analyses, the added value of the dimensions over DSM-IV categories should be further investigated.

Dimensions should be shown to have potential added value on top of traditional psychopathology measures. They could be used as more specific phenotypes in biological etiological research to overcome the heterogeneity and comorbidity that has hampered research with DSM-defined research groups. In addition, dimensions could be used as more specific predictors of disease-course and treatment response. The applicability of dimensions for these purposes still needs to be thoroughly investigated, but they could be promising leads to improving diagnostics and the specificity of treatment indications.

In conclusion, we present an integrated six-dimensional model to assess different symptoms of depression and anxiety that does justice to the heterogeneity of anxiety and consists of easily measurable dimensions. These dimensions could eventually be used as more specific phenotypes in etiological research and to describe patients’ symptom patterns in clinical settings.