

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/22851> holds various files of this Leiden University dissertation

**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 4**

### **THE TRIPARTITE MODEL FOR ASSESSING SYMPTOMS OF ANXIETY AND DEPRESSION: PSYCHOMETRICS OF THE DUTCH VERSION OF THE MOOD AND ANXIETY SYMPTOM QUESTIONNAIRE**

Behaviour Research and Therapy 2007; 45, 1609-1617

Edwin de Beurs  
Margien E. den Hollander-Gijsman  
Stephanie Helmich  
Frans G. Zitman

## **Abstract**

### ***Aim***

The tripartite model conceptualizes symptoms of depression and anxiety in three groups: low positive affect and anhedonia, which is specific to depression, somatic arousal, which is unique to anxiety, and nonspecific general distress. The Mood and Anxiety Symptom Questionnaire (MASQ) was developed to measure these symptom domains. This study reports on the psychometric properties of the Dutch translation of the MASQ.

### ***Method***

The questionnaire was completed by a population-based sample and by patients with anxiety and/or mood disorders. Scores of these respondent groups were compared to assess the discriminant validity of the MASQ and evaluate the appropriateness of the tripartite model.

### ***Results***

The psychometric properties of the translated MASQ were highly satisfactory. In accordance with the model, we found the MASQ to comprise three main scales, which discriminate well between subgroups of patients with mood and anxiety disorders.

### ***Discussion***

Overall, like the English version the Dutch translation of the instrument appears to be a reliable and valid measure of symptoms of depression and anxiety, conceptualized as comprising three groups of symptoms. The Dutch MASQ is better able to distinguish unique aspects of mood and anxiety disorders than other self-report instruments.

#### 4.1 Introduction

Whether a valid distinction can be made between anxiety and depression is subject to much debate. Both disorders show considerable overlap in symptomatology, making it sometimes hard to decide which diagnosis best fits the clinical picture (e.g. generalized anxiety disorder vs. major depression). Consequently, instruments assessing the key symptoms of mood and anxiety disorders show high convergence, partly due to the similarity in item content of such measures. This is unfortunate and hinders progress in research into the shared and distinct features of both disorders. For instance, investigating whether mood and anxiety disorders have a different etiology or a different biological background is hampered when the difficulty in psychometrically distinguishing both disorders is not resolved. After a comprehensive review of the existing literature on the relation between anxiety and depression, Clark and Watson proposed the tripartite model for depression and anxiety (Clark & Watson, 1991). The model proposes one general distress factor and two additional factors specific to anxiety and depression. The model is based on Clark and Watson's and Tellegen's earlier work (Tellegen, Watson, & Clark, 1999) on dissecting mood into two independent components: negative affect (NA) and positive affect (PA). NA is characterized by aversive emotional states, such as feeling upset, angry, guilty, afraid, sad, scornful, and disgusted. PA represents positive emotional states such as feeling active, delighted, interested, and enthusiastic. Lack of PA is best described by terms such as feeling tired, sluggish, feeling that nothing is enjoyable, and not having fun in life (Clark & Watson, 1991). Not only are both mood dimensions fairly independent as evidenced by moderate to low intercorrelations, their separateness is also supported by distinctive correlational patterns with other variables, such as social activity (only with PA) and health complaints (only with NA). Furthermore, personality trait characteristics such as neuroticism are more strongly associated with NA, whereas extraversion has a stronger association with PA. In addition to PA and NA, Clark and Watson proposed a third dimension, physiological hyperarousal, which encompasses symptoms such as tenseness, shortness of breath, feeling dizzy or lightheaded, trembling and shaking. This dimension has also been labeled in the literature as somatic arousal (SA). These symptoms appear to be better in differentiating between anxious and depressed patients than symptoms reflecting anxious mood per se (Clark & Watson, 1991). The model explains the high concurrence of anxiety and depression by proposing that both disorders share the dimension of NA. Unique to mood disorders is a lack of PA; unique to anxiety disorders (especially panic disorder) is physiological hyperarousal. The tripartite model for depression and anxiety has found broad acceptance, not only with adult patients (Marshall, Sherbourne, Meredith, Camp, & Hays, 2003; Joiner et al., 1999; Keogh & Reidy, 2000) but also in child psychiatry populations (Chorpita

& Daleiden, 2002).

Watson and Clark (1991) developed a 90-item self-report questionnaire to measure the three dimensions of the tripartite model, the Mood and Anxiety Symptom Questionnaire (MASQ). The MASQ contains two scales with symptoms specific to depression and anxiety. The first scale, Anhedonic Depression (AD) is meant to measure (lack of) PA. The second scale, Anxious Arousal (AA) measures symptoms of SA. The remaining items are all relatively non-specific and measure NA or general distress (GD). However, based on their content, these items are further subdivided into a third, fourth and fifth scale containing depression General Distress Depression (GDD), General Distress Anxiety (GDA) and General Distress Mixed (GDM) symptoms. Research findings regarding the validity of these scales are favorable for the MASQ (Reidy & Keogh, 1997; Watson et al., 1995), but the dimensional structure of the MASQ as comprising five scales is not clearly supported. Investigations into the factors of the MASQ with clinical and normal samples have generally found the MASQ to comprise three scales (Bedford, 1997; Reidy & Keogh, 1997; Watson et al., 1995), which is actually more in accordance with the tripartite model.

The present study set out to translate the MASQ in Dutch and evaluate the psychometric characteristics of the Dutch translation in a large sample of psychiatric outpatients with mood and anxiety disorders and a representative sample of the general population. The MASQ was translated according to the guidelines of Widenfelt and colleagues (Van Widenfelt, Treffers, De Beurs, Siebelink, & Koudijs, 2005). First, we investigated whether the dimensional structure of the MASQ was preserved in the Dutch translation with exploratory factor-analysis. We compared the factor structure for the translated MASQ with published results from US (Watson et al., 1995) and British (Keogh & Reidy, 2000) samples. Next, we evaluated the psychometric characteristics of the translation by assessing indices of reliability (internal consistency). Concurrent and divergent validity was assessed by comparing the MASQ with other self-report questionnaires or rating scales (Campbell & Fiske, 1959). The discriminant validity of the MASQ scales was evaluated by comparing scores of psychiatric outpatients with the population-based sample. Also we compared scores of subgroups of patients with specific diagnoses. With these latter analyses, we could investigate the uniqueness of the PA and somatic anxiety subscales of the MASQ for depression and anxiety, respectively. We hypothesize lower PA scores for patients with depressive disorders and higher somatic anxiety scores for patients with anxiety disorders, especially patients with panic disorder.

## **Method**

### ***Sample and procedure***

A patient sample was composed of 950 patients from three outpatient clinics of the Rivierduinen Psychiatric Hospital (675 consecutive admissions from Leiden; 158 from Alphen a/d Rijn, and 117 from Voorhout). All patients were referred to these clinics by their General Practitioner for a mood, anxiety or somatoform disorder. The sample contained 625 (65.8%) females; the average age was 36.2 years,  $sd=11.6$ , range 17–68). The diagnosis was assessed with a standardized diagnostic interview, the Mini International Neuropsychiatric Interview (MINI-plus; (Sheehan et al., 1998), which was carried out by a research nurse (a psychiatric nurse or a psychologist). In the assessment session with the research nurse self-report questionnaires were administered through a computer program and the research nurse completed several rating scales. The entire assessment session took about 120 min. Total of 894 patients met criteria for one or more DSM-IV diagnosis (94.9% of the sample); 261 (27.5%) had one diagnosis, 286 (30.1%) had two, 189 (19.9%) had three, 80 (8.4%) had four, and 78 (8.2%) had more than four diagnoses. In the sample, 498 anxiety disorders were diagnosed and 490 mood disorders, 270 of these were comorbid cases.

A sample of 200 respondents from the general population was obtained by randomly picking names from a listing in the phonebook of Leiden and vicinity. Special care was taken to ensure that the sample was similar to the general population on relevant variables such as size of the place of residence and gender (two-staged proportioned stratified sampling (Moser & Kalton, 1979). Various techniques were employed to optimize the response rate (Dillmann, 1978), such as telephoning potential respondents for consent before sending them questionnaires, inclusion of a cover letter in which the importance and the scientific purpose of the study was underlined, and sending a follow-up letter to those, who had not returned the questionnaire within 3 weeks. Thus, 363 persons were approached and invited to partake in a study “investigating questionnaires for the assessment of emotional functioning” of which 255 (70%) agreed to participate. A total of 204 questionnaires were returned, of which 200 contained usable data (78% of the questionnaires that had been sent out and 55% of all contacted potential respondents). A response rate of 55% is substantial for a mail survey, boosting our confidence in the sample as being representative. We compared demographic characteristics of the sample (gender, age, marital status, education level, and religiosity) with the general population. This indicated that there was no sample bias, except for a slight under representation of the age group 18–25 and an overrepresentation of respondents aged 65 and older. This was probably due to the fact that younger people are less likely to be listed in the telephone book (our first source of respondents) because they use nowadays predominantly mobile phones in the Netherlands. Fifty-five percent of the respondents were female; the mean

age was 47.5 years (sd=15.0, range=18–88); 69.0% was married; 52.0% held a fulltime or part-time job; 25.5% were stay at home wives or mothers.

### ***Measures***

#### ***MASQ***

The MASQ (Watson & Clark, 1991) contains “a list of feelings, sensations, problems and experiences that people sometime have” (instructions to the respondent). The respondent is asked to indicate on a Likert-scale (0=not at all, 4=extremely) how much they have felt or experienced these feelings or thoughts in the past week including today. Watson and Clark grouped 77 of the 90 items of the MASQ in 5 subscales based on their content. Three subscales measure relatively nonspecific symptoms of general distress. Due to their similarity with DSM-III-R criteria for mood or anxiety disorders, items were assigned to either the GDM, GDA, or GDD subscale. Furthermore, two subscales comprise symptoms specific to anxiety and depression. Seventeen items reflecting symptoms of somatic tension and hyperarousal were grouped in the AA scale. Eight loss of interest items and 14 PA items composed the AD scale.

Three independent translations of the MASQ were made by native Dutch researchers with ample experience in translation of measurement instruments (A.M. van Hemert, M.D., Ph.D., J. Goekoop, M.D., Ph.D., and E. de Beurs, Ph.D.). The three translations were compared and discrepancies in the translations were discussed until consensus on a final translation was reached. Next, a native speaker (B.M. van Widenfelt, Ph.D.) translated this version back into English. The original questionnaire and the back translation were compared and where discrepancies were found minor revisions were applied to the translation. For 12 of the 90 items, minor revisions in phrasing were deemed necessary. These revised items were discussed among the original translators and again a back translation was performed with a satisfactory outcome.

#### ***Other instruments***

All patients and the respondents from the population sample completed the Brief Symptom Inventory (BSI, (Derogatis, 1975). On this checklist of 53 symptoms, the respondent indicates to what extend they have been bothered by each symptom in the last week, including today (0=“not at all”, 4=“extremely”). The BSI comprises among others subscales for somatic complaints, depression, anxiety, phobic avoidance and interpersonal sensitivity. The total score on the BSI is generally perceived as a highly reliable index of general psychopathology. Patients with a current major depression or dysthymia completed the Beck Depression Inventory II (BDI; (Beck & Steer, 1987).

Diagnostic status was assessed with the MINI-plus (Sheehan et al., 1998). The MINI-plus is a standardized diagnostic interview comprising 23 modules

in which the presence or absence of DSM criteria for the main psychiatric disorders (mood, anxiety, psychotic, somatoform, and eating disorders) is investigated. Each module starts with one or two screening questions. If these are answered affirmatively, additional questions from the module are asked. Lecrubier and colleagues (1997) report sufficient reliability for most modules. Inter-rater reliability ranged from  $k=0.88$  to  $1.00$ , test-retest reliability ranged from  $0.76$  to  $0.93$ , validity was demonstrated by sufficient concordance with the CIDI ( $k$  ranged from  $0.36$  for generalized anxiety disorder to  $0.82$  for alcohol dependence).

In addition, the psychopathology of the patients was rated by the research nurse on a shortened version of the Comprehensive Psychiatric Rating Scale (CPRS) comprising 25 items in three subscales, the Montgomery Asberg Depression Rating scale (MADRS (10 items), the Brief Anxiety Scale (10 items) and the Retardation Scale (5 items) (Goekoop et al., 1991). Items on the CPRS (e.g., "pessimistic thoughts", "worries about minor issues") are rated on a 7-point scale anchored at 4 points (1, 3, 5, and 7) with different response options for each item. The research nurse completed the Global Assessment of Functioning scale of the DSM, a scale for impairment in functioning due to the psychiatric complaints ranging from 0 to 100 (American Psychiatric Association, 1994), and the Clinical Global Impression (CGI; (Guy, 1976) (severity of illness scored on a 7 point scale ranging from 1="normal, no complaints" to 7="extremely ill"). Research nurses were extensively trained in administration of the rating scales and in the diagnostic interview. Each new research nurse followed an intensive 2-week training with an experienced nurse in performing the assessments, before being allowed to do ratings on her own. In addition, biweekly training sessions of 2 h were organized continuously in which invited speakers taught about psychiatric disorders and videotaped patients were conjointly rated by the group of research nurses to improve interrater reliability. For a small subset of patients ( $n=44$ ) the assessment session was audio taped. After listening to these tapes another research nurse rated patients again and this revealed sufficient interrater reliability (average concordance between raters was sufficient. Average Cohen's  $k=0.60$  for the CPRS (average  $k=0.59$  for 19 interview items and average  $k=0.63$  for 6 observational items),  $k=0.73$  for the GAF-score (recoded into 5 categories), and  $k=0.55$  for the CGI-score).

### ***Statistical analysis***

First, the frequency distributions of scores on the translated items were investigated (mean, sd, skewedness, and kurtosis). Next, the factor structure of the instrument was investigated with exploratory factor analysis, utilizing parallel analysis to decide on the number of factors to retain (O'Connor, 2000). The rotated factor solution was compared with published results of US and British samples. Reliability was investigated by assessing the internal



consistency of scales. Validity was evaluated by assessing the convergence with parallel tests (bivariate correlations). Finally, we assessed the ability of the MASQ scales to differentiate between the patient sample and the population sample, as well as between diagnostic subgroups within the patient sample with t-test.

### **4.3 Results**

#### ***Basic psychometrics and construct validity (factor structure)***

Inspection of the frequency distributions of the individual items of the translated questionnaire did not reveal substantial deviation from normality, implying no need to alter phrasing of any items. For some items, scores from the population-based sample were skewed, but this is understandable given the low prevalence of certain feelings in the general population (e.g. "thought about death or suicide").

To investigate the factor structure the 90 items of the MASQ were subjected to an exploratory factor analysis. Parallel analysis suggested retaining three factors and the screening occurred after the third or fourth factor. Thus, a three-factor solution was chosen. Next, factor loadings were inspected to allocate items to subscales. Utilizing two criteria of a primary loading  $>0.30$  and sufficient purity (a cross-loading  $<0.20$ ), the first factor (NA) comprised 20 items; items 4, 6, 8, 13, 16, 17, 20, 22, 24, 26, 28, 29, 42, 47, 53, 64, 74, 77, 84, 89), the second factor (PA) comprised 22 items (1, 11, 14, 18, 23, 27, 30, 35, 36, 38, 40, 41, 43, 46, 49, 54, 58, 62, 68, 72, 78, 86), and the third factor (somatic anxiety or SA) 18 items (9, 25, 45, 48, 52, 55, 57, 61, 63, 65, 67, 69, 73, 75, 79, 81, 87, 88). The factor loading were quite similar to the results of Watson, Clark, et al. (1995) and the factor solution was in almost perfect agreement with the results of Keogh and Reidy (2000). As can be seen in Table 4.1, the best correspondence is found for the PA scale. Our items match almost perfectly with the Keogh and Reidy (2000) solution and the Watson, Clark, et al. (1995) solution. Regarding the SA dimension, the match in items is still substantial: 16 of 18 items of the present solution match up with Watson et al., 15 with Keogh and Reidy. Finally, comparison of the item composition of the NA scale in the three samples again reveals substantial overlap: 17 of 20 items match with the Watson et al. solution, 14 of 20 match with the Keogh and Reidy solution.

#### ***Reliability of the scales***

The reliability indices of the scales (internal consistency) and intercorrelations among the scales are presented in Table 4.2. Reliability of the three scales was excellent: all  $\alpha \geq 0.88$ . The correlation between the NA and the PA scale was substantial ( $r=0.62$ ), but the SA scale correlated only moderately with the NA scale ( $r=0.53$ ) and low with the PA scale ( $r=0.35$ ), indicating a shared variance of 28% and 12%, respectively. Correlations among the scales were generally

		Watson et al. (1995) <sup>a</sup>				Keogh & Reidy (2000)				
		GD	A-PA	SA	Not assig.	NA	PA	SA	Not assig.	Total
Present sample	NA	17			3	14			6	20
	PA		21		1		22			22
	SA			16	2			15	3	18
	Not assigned	5	1	2	22	7	1	1	21	30
Total		22	22	18	28	21	23	16	30	90

NA = negative affect, PA = positive affect, SA = somatic anxiety.  
<sup>a</sup> In Watson et al. (1995) Table 6 the three factors found are named General Distress – GD, Anhedonia/Positive Affect – A-PA, and Somatic Anxiety-SA)

Table 4.1 Number of corresponding items in various factor solutions.

	NA	PA	SA
No. of items:	20	22	18
NA	(0.96)		
PA	0.62	(0.96)	
SA	0.53	0.35	(0.91)

Note: Scale reliabilities are shown between parentheses. All correlations are significant at the 0.001 level (2-tailed).  
 NA = Negative Affect, PA = Positive Affect, SA = Somatic Anxiety

Table 4.2 Reliability and correlation coefficients between MASQ scales and reliabilities (Cronbachs *a*).

	N	NA	PA	SA
<i>Rating scales:</i>				
GAF	596*	-0.32	-0.33	-0.34
CGI	599*	0.30	0.31	0.28
MADRS	935	0.69	0.64	0.50
BAS	848	0.50	0.42	0.57
INH	515	0.44	0.46	0.27
<i>Self-report:</i>				
BDI-II	583**	0.80	0.61	0.47
BSI -dep	929	0.86	0.63	0.43
BSI -anx	929	0.62	0.37	0.62
BSI -pho	929	0.53	0.34	0.47
BSI -som	927	0.45	0.32	0.84
BSI -int	929	0.69	0.41	0.33
BSI -tot	929	0.82	0.52	0.64

NA = Negative Affect, PA = Positive Affect, SA = Somatic Anxiety; GAF = Global Assessment of Functioning; CGI = Clinical Global Impression, MADRS = Montgomery Asberg Depression Rating Scale, BAS = Brief Anxiety Scale, INH = Inhibition; BDI = Beck Depression Inventory Revised; BSI = Brief Symptom scale; dep = depression, anx = anxiety, pho = phobic anxiety, som = somatic complaints, int = interpersonal sensitivity, tot = Total score (all correlations  $p < .001$ )

\* Data are available for less patients since these measures were later introduced in the assessment battery.

\*\* The BDI-II was administered only if patients met criteria for a mood disorder.

Table 4.3 Correlation of the MASQ scales with rating scales and self-report measures.

lower than the correlation between scales composed according to the allocation of Watson, Clark, et al. (1995) would be. With calculations based on our data the correlation between the AD and AA scales would amount to  $r=0.49$ ; between the PA–SA scales the association is  $r=0.35$ , a decrease from 24% to 12% shared variance. The current two scales are clearly more distinct.

Table 4.3 presents the correlation coefficients between the three MASQ scales and other measures of psychopathology. Both the GAF and the CGI scores show modest and roughly equal correlations with the MASQ scales. In contrast the MADRS, a rating scale for depression shows the highest convergence with

	Patients (950)		Population (200)		T (1148)	Cohen's <i>d</i>
	mean	sd	mean	sd		
NA	2.57	0.97	1.44	0.50	16.20*	1.46
PA	4.00	0.77	3.25	0.77	12.55*	0.97
SA	1.88	0.74	1.25	0.36	11.80*	1.08

NA = negative affect, PA = positive affect, SA = somatic anxiety.  
\*  $p < 0.001$

Table 4.4 Comparison of mean scores on five MASQ scales of the patient and population sample.

	No depression (N = 460)		Depression (N = 490)		t(948)	d	No anxiety (N = 452)		Anxiety (N = 498)		t(948)	d
	mean	sd	mean	sd			mean	sd	mean	sd		
NA	2.09	0.79	3.03	0.90	16.98*	1.11	2.34	0.92	2.79	0.96	7.27*	0.48
PA	3.61	0.75	4.37	0.58	17.53*	1.13	3.91	0.81	4.09	0.72	3.57*	0.23
SA	1.67	0.64	2.09	0.77	6.25*	0.59	1.71	0.62	2.04	0.80	7.14*	0.46

NA = negative affect, PA = positive affect, SA = somatic anxiety.  
\*  $p < .001$ .

Table 4.5 Mean scores (and sd's), results of t-tests and effectsize of the difference when analysing two contrast: patient with and without a mood disorder and patients with and without an anxiety disorder.

the NA and the PA scale ( $r=0.69$  and  $0.64$ , respectively), whereas the BAS (rating of anxiety) has stronger correlations with the SA scale than with the PA and NA scales. A similar pattern of correlation emerges with the self-report measures: The BDI-II and the BSI-dep scale correlate most strongly with the NA and PA scales. The high correlation between the SA and the BSI Somatic complaints subscale ( $r=0.83$ ) reflect the predominance of somatic markers of anxiety in the SA scale (Table 4.3).

### ***Discriminant validity***

A first test for the criterion related validity of the instrument is its ability to discriminate between patients and the normal population. We compared both groups with t-tests. The means, sd's, results of the t-tests, and the effect size of the difference (Cohen's d) are listed in Table 4.4. All scales discriminate well between patients and normal controls and statistical significance is upheld after Bonferonni correction for multiple testing. Differences among the various subscales in discriminant validity are small.

Demonstrating that the MASQ scales are able to discriminate between patients and respondent from the general population may be useful for certain research goals (e.g. screening in epidemiological research), but a test of the validity of the MASQ should also encompass assessment of the ability of the instrument to discriminate between groups of patients, especially patients who suffer predominantly from anxiety vs depressed patients. The ability of the PA scale to discriminate between patients with and without a mood disorder (as PA is supposedly unique to depression) and the ability of the SA scale to discriminate between patients with and without an anxiety disorder are especially relevant for the MASQ. Therefore, we compared MASQ scale scores for different subgroups of patients in our patient sample.

Based on the DSM diagnosis according to the MINI we selected from the patient sample several subgroups: patients with and without a current diagnosis of mood and patients with and without an anxiety disorder. Table 4.5 presents mean scores of the subgroups of patients and results of the comparison with t-tests. The largest difference between depressed and non-depressed patients is on the PA scale, closely followed by the NA scale. The SA scale is less suited to distinguish depressed from non-depressed cases. This finding supports the validity of the PA scale. Regarding the anxiety contrast, the results are somewhat less favorable for the measure. The NA and SA scales appear to be the best in differentiating between patient with and without an anxiety disorder. However, the difference between both groups is not larger on the SA scale as compared to the NA scale. This finding does not support the presumed uniqueness of the SA scale for anxiety.

### **4.4 Discussion**

Until now, most of the psychometric research with the MASQ has been done with non-clinical samples (usually undergraduate students) or with relatively small patient samples. We administered the questionnaire to a large patient group with the relevant disorders: mood and anxiety. Furthermore, the diagnostic status of these patients was comprehensively assessed in a diagnostic interview by well-trained research nurses. Data on the diagnostic status of the respondents enabled us to investigate the discriminant validity of the MASQ by comparing scores on subscales from distinct clinical subgroups.

Administration of other, well established selfreport measures enabled us to investigate convergent and divergent validity.

First of all, the present findings suggest that the MASQ has been adequately translated for use in the Netherlands: The items show satisfactory psychometric properties and, although the factor solution differs considerably from the five subscales originally proposed by Watson, Clark, et al. (1995), the solution is in accordance with the results of factor analyses from US and English datasets. The three scales of the instrument are reliable considering their high internal consistency coefficients. The validity of the scales of the instrument is also supported by substantial correlations with other instruments. The MASQ-scales have sufficient discriminant power. In sum, the validity coefficients favor three subscales for the MASQ, rather than the original conceptualization of the instrument in five subscales.

The factor structure of the translated MASQ was concordant with results obtained by Keogh and Reidy (2000), but less so with results of Watson, Clark, et al. (1995). Discongruity can stem from two sources: crosscultural differences or problems with the translation of the MASQ into Dutch. Both effects are difficult to disentangle, but a first attempt could be to compare all three-factor solutions amongst each other. This comparison revealed the best concordance between the Dutch results and the results based on the British sample, suggesting a cross-cultural difference rather than a difference due to a problematic translation of the instrument.

The findings of the factor analyses support the tripartite structure of depression and anxiety with three distinct factors. Moreover, a three-factor solution has been repeatedly suggested in the literature as best fitting the data (Bedford, 1997; Reidy & Keogh, 1997) and is in accordance with the formulation of the tripartite model. However, two findings deserve more critical consideration than they have been given in the previous studies. The PA factor comprises 22 of the 24 reversed keyed items of the MASQ. This result is not due to the translation into Dutch, but replicates the results of Watson, Clark, et al. (1995) and Keogh and Reidy (2000). Previous studies fail to comment on this potential flaw of the instrument. The grouping of all reversed keyed items in one factor is an unfortunate outcome as it suggests the possibility of a method effect underlying this factor, rather than a true distinct construct. For the factor structure of the MASQ Watson, Clark, et al. (1995) predicted three broad factors with one factor being a "specific depression factor that is on one end defined by items reflecting energy, enthusiasm and high PA and on the other by items reflecting anhedonia, loss of interest, and low PA"(p. 16). Consequently, they grouped the "lack of interest" items under the AD scale. The results of exploratory factor analysis of the present study, as well as results from the study by Keogh and Reidy (2000) and Watson, Clark, et al. (1995) themselves do not support such an item allocation. Nitschke and colleagues

tested the homogeneity of the AD scale with confirmatory factor analysis and also concluded that this scale comprised two separate constructs: the eight "lack of interest" items and the 14 reversed scored items describing positive feelings. According to the present findings, the lack of interest items belong to the NA dimension and are thus no longer specific for mood disorders (Nitschke et al., 2001). The PA dimension now only comprises items describing positive feelings. Future revisions of the MASQ should encompass items belonging to lack of PA that describe negative feelings.

Further research is needed to investigate the validity of the dimensions of the tripartite model and the ability of the MASQ to adequately assess these. Strengths of the present study are the use of a representative population sample and the large dataset of patients. A limitation of the study is that investigation of concurrent and divergent validity was restricted to comparison with other self-report scales and ratings by an observer. Furthermore, discriminant validity was assessed by comparison of diagnostic subgroups. Both approaches have their drawbacks. First, the other scales used to validate the MASQ have their own flaws and weaknesses. Secondly, forming diagnostic subgroups based on a diagnostic interview such as the MINI will never be perfect and some patients will have been misclassified. Therefore, additional validation by other means is called for. For instance, comparison of MASQ-scores with neuroendocrinological or neuro-imaging data or outcomes of neuro-psychological testing will yield valuable data regarding the validity of the tripartite model (Shankman & Klein, 2003). Presently we are evaluating a shortened scale comprising 10 items for each of the three concepts of the tripartite model. The scale is included in a large longitudinal study (Netherlands Study on Depression and Anxiety, NESDA). The predictive validity of these shortened scales on the long-term course of mood and anxiety disorders will be investigated. Thus, the ability of the MASQ to assess changes in symptomatology and to predict the course of anxiety and depression symptoms over time will be investigated. Until now, research on the MASQ has been limited to cross-sectional data. Its sensitivity to change over time or to treatment effect has not been established. Testing of etiological models for depression and anxiety with the MASQ as dependent variable may shed more light on the validity of the measure and the value of the tripartite model (De Beurs et al., 2005).





