OUT OF THE BOX

Moving from categories to dimensions in the phenomenology of depression and anxiety

Margien E. den Hollander-Gijsman

OUT OF THE BOX

Moving from categories to dimensions in the phenomenology of depression and anxiety

Cover illustration: M.C. Escher's "Man with Cube"

Print: GVO drukkers & vormgevers B.V. | Ponsen & Looijen
Lay-out: Celines Zaak

OUT OF THE BOX

Moving from categories to dimensions in the phenomenology of depression and anxiety

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 11 december 2013
klokke 13:45 uur

door

Margien Elisabeth den Hollander-Gijsman
geboren te Leiden
in 1970

Promotiecommissie

Promotores:	Prof. dr. F.G. Zitman Prof. dr. N.J.A. van der Wee
Copromotor:	Dr. E. de Beurs
Overige leden:	Prof. dr. A.J.L.M. van Balkom (VUMC) Prof. dr. A.J.W. van der Does Prof. dr. R.C. van der Mast Prof. dr. R.J.J.M. Vermeiren

Table of Contents

Chapter 1	Introduction	7
Chapter 2	Routine Outcome Monitoring in the Netherlands: Practical Experiences with a Web-Based Strategy for the Assessment of Treatment Outcome in Clinical Practice	21
Chapter 3	Comorbid Depression and Anxiety: symptom and functional severity in the clinical setting	39
Chapter 4	The tripartite model for assessing symptoms of anxiety and depression: Psychometrics of the Dutch version of the mood and anxiety symptom questionnaire	53
Chapter 5	Distinguishing between depression and anxiety: a proposal for an extension of the tripartite model	69
Chapter 6	Distinguishing symptom dimensions of depression and anxiety: An integrative approach	91
Chapter 7	Summary, Discussion and Future perspectives	111

Reference List	127
Samenvatting	141
Curriculum Vitae	149
List of Publications	151
Dankwoord	155

CHAPTER 1

INTRODUCTION

This thesis focuses on the phenomenology of depression and anxiety disorders, and pleads for a transition from a categorical to a multidimensional approach in the description of these disorders.

The relationship between depression and anxiety disorders has drawn much attention for several decades. This continuing interest is mainly due to the high comorbidity between the disorders, and the similarities between effective treatments. In line with that, there is an on-going discussion about whether depression and anxiety disorders should be considered as separate diagnostic entities with specific phenomenological and neurobiological characterizations, or should be considered as manifestations of the same disorder.

The way the phenomenology of these disorders is described influences the design and the results of studies on the cause and treatment of the disorders. Although the development of clear categorical diagnostic criteria, as described in the widely used diagnostic system of the DSM-IV (Diagnostic and statistical manual of mental disorders (American Psychiatric Association, 1994, p. 111)), has advanced the field considerably, it has become more and more clear that a mere categorical approach hampers further progress. In this dissertation we set out to make a step forward from the use of categorical conceptualization to a dimensional approach of depression and anxiety.

In this introduction we first present the advantages and disadvantages of categorical systems. Second, we describe the general benefits of a dimensional approach, and present several dimensional approaches that have been proposed in the literature. Third, we discuss the tripartite model (a well-known dimensional approach) at length, and present some of the criticism on this model. Fourth, we explain how the tripartite model can be improved. Finally, we describe the main purpose and outline of this thesis.

1.1 Categorical systems

Categorical systems: advantages

In psychiatry the DSM-IV and ICD-10 (International Statistical Classification of Diseases and Related Health Problems) are the commonly used categorical sets of criteria for psychiatric diagnoses. Each disorder is defined by its symptoms and criteria are given which have to be met, for a diagnosis to apply. As an example the DSM-IV criteria for a Major Depressive Episode are shown in textbox 1.1.

The development of clear diagnostic criteria has advanced the field considerably. Since the introduction of the DSM and ICD, communication among professionals and among researchers about disorders and individual patients has become more straightforward. On top of that, comparing research results (e.g. in reviews and meta-analyses) has become much more effective and accurate.

DSM-IV criteria for Major Depressive Episode

Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood, or (2) loss of interest or pleasure.

1. Depressed mood (either subjective report or observation)
2. Markedly diminished interest or pleasure in activities
3. Significant weight loss when not dieting or weight gain
4. Insomnia or hypersomnia
5. Psychomotor agitation or retardation nearly every day
6. Fatigue or loss of energy
7. Feelings of worthlessness or excessive or inappropriate guilt.
8. Diminished ability to think or concentrate, or indecisiveness.
9. Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or specific plan.

Textbox 1.1

Categorical systems: drawbacks

As a basis for research into psychopathology (e.g. endophenotype, genotype, trauma's, personality), a categorical approach has serious drawbacks. One problem with a categorical diagnostic system is the arbitrary nature of its diagnostic boundaries. For instance, the presence or absence of just one symptom can make the difference between meeting and not meeting the diagnostic criteria for a disorder. Indeed, research into etiological or risk factors for mental disorders shows that subsyndromal cases are very similar to cases fully meeting the diagnostic criteria for the disorder (McGorry, Hickie, Yung, Pantelis, & Jackson, 2006; De Beurs et al., 1999).

Another problem with categorical diagnostic classification is that patients with different symptom profiles may nevertheless meet criteria for the same diagnosis, which results in heterogenic groups of patients within a diagnostic class. To illustrate the heterogeneity of patients with a depression, we describe and compare the psychopathology of two patients in textbox 1.2. Both patients meet diagnostic criteria for a Major Depressive Disorder (MDD) according to the DSM-IV.

Although both patients fulfil the criteria of a MDD, they display a very different pattern of symptoms. This is possible because in the DSM-IV fulfilling five out of nine criteria described for MDD (see textbox 1.1), is enough to be eligible for a diagnosis. On top of that, several criteria contain both the increase and the

Two patients with Major Depressive Disorder (MDD)

Patient A (female, 36 years of age) is feeling sad and worthless, has problems with falling asleep, and has lost weight. She is quite agitated, is feeling very guilty towards her family (husband and two young children), and often wished she would not wake up the next morning ("that would be better for everyone").

Patient B (female, 29 years of age) does not enjoy her life as much as she used to do. She feels bogged down and sluggish. She sleeps a lot and has energy loss, gained 5 kilograms in the past 3 months and does not get anything done and can't concentrate.

Textbox 1.2

decrease of certain behaviour (e.g. weight loss OR weight gain). This illustrates that a group of patients with MDD may indeed be very heterogeneous group.

To complicate matters further, in the current diagnostic systems it is quite likely that a psychiatric patient meets criteria for two or even more, psychiatric disorders simultaneously. In fact, the majority of patients meet criteria of more than one psychiatric disorder concurrently (e.g., Brown, Campbell, Lehman, Grisham, & Mancill, 2001; De Graaf, Bijl, Smit, Vollebergh, & Spijker, 2002; Kessler et al., 1996). This is certainly true for mood and anxiety disorders (e.g., Belzer & Schneier, 2004). High rates of comorbidity between anxiety disorders and depression have been reported in the general population (Kessler et al., 1996), in primary care (Roca et al., 2009) and in secondary care (Brown et al., 2001).

As a syndrome is defined as a permanent combination of symptoms, it can be concluded that DSM-IV diagnoses are not real syndromes. Rather, they are more or less arbitrarily constructed clusters of symptoms. The etiology and pathophysiology of depression and anxiety disorders are unknown, and illness-specific treatments are absent. For instance, a wide range of emotional disorders respond similarly to the same psychosocial and psychotropic drug treatment (Brown & Leyfer, 2009). In this respect, the growing number of disorders in each new version of the DSM has led to questions about the discriminant validity of the categories. For example, the DSM-II had only three anxiety disorders (American Psychiatric Association, 1968) and the more recent DSM-IV includes twelve anxiety disorders (American Psychiatric Association, 1994). The distinction between these various anxiety disorders is predominantly based on their symptomatic phenomenology and not on established knowledge

regarding their etiology or treatment response to different (pharmacological or psychotherapeutical) agents.

1.2 Multidimensional models

General benefits of a dimensional approach

Many researchers and clinicians agree that the fields of psychiatry and clinical psychology could benefit from a dimensional model of the psychopathology of common mental disorders (Clark, 2005; Cuthbert, 2005; Krueger, Watson, & Barlow, 2005). A multidimensional model provides each patient with a symptom profile on a number of dimensions of psychopathology, rather than a dichotomous presence or absence of one or more disorders. It offers a more specific and accurate description of individuals, covers the full range of severity from healthy to severe psychopathology and may help to avoid stereotyping. Moreover, a multidimensional approach has the potential to associate psychopathology more accurately to etiological factors and biological markers and may predict more accurately which patients will benefit from a specific treatment (Kupfer, 2005; Widiger & Samuel, 2005).

The idea that a dimensional model has great benefits over a categorical system has gained much support in the last decades. Several dimensional models of depression and anxiety have been proposed. However, there is yet no clarity about which model is to be preferred. Among the models proposed so far, a distinction can be made between 'disorder-based models' and 'symptom-based models'.

Disorder-based models

There are several disorder-based approaches. The first line of research focuses on finding dimensions with optimal ability to discriminate between DSM patient groups, for example between generalized anxiety disorder and depression (Kessler et al., 2002; Wittchen et al., 2002). In this approach, the goal is to determine the symptom group(s) that represent(s) the most characteristic features of a DSM diagnosis.

A second disorder-based approach that is often referred to as a dimensional approach is the clustering of DSM-categories in a hierarchical way, with the aim to find common and distinctive features among various disorders based on their co-occurrence. In two large studies in respectively the United States (National Comorbidity Survey) and The Netherlands (NEMESIS study) the latent structure of common mental disorders was studied in the general population (Krueger, 1999; Vollebergh et al., 2001). In both studies the results of confirmatory factor analysis show an optimal fit for a hierarchical three-factor model, in which the internalizing problems are represented by two latent factors: 1) anxious misery (major depression, dysthymia and generalized anxiety disorders), and 2) fear (the other anxiety disorders). The third factor in these two studies

represents respectively the external disorders (Krueger, 1999), and alcohol/drug dependence (Vollebergh et al., 2001). For future research on common mental disorders both studies suggest to focus on core processes and core psychological features of the disorders, rather than on their manifestations as distinguished disorders, and on further differentiating between subtypes of disorders.

Referring to the studies of Krueger and Vollebergh, Watson states that "there are now sufficient data to eliminate this rational system and replace it with an empirically based structure that reflects the actual similarities among disorders" (2005). With 'this rational system' Watson means the DSM-IV, and he subsequently makes a suggestion for the structure of emotional disorders in the next version of the DSM, the DSM-5. He suggests to use a quantitative hierarchical model in which the mood and anxiety disorders are taken together to form an overarching class of emotional disorders with 3 subclasses (distress disorders, fear disorders and bipolar disorders) (Watson, 2005).

Although practical and intuitive, a disadvantage of this disorder-based approach is that the 'dimensions' are broad latent variables of existing DSM-IV categories, which are arbitrary and heterogeneous as discussed above. This means that the problems with a categorical system (arbitrary boundaries, comorbidity and heterogeneity of patient groups) are not fully resolved and seep into the new system.

Symptom-based models

A symptom-based dimensional approach is characterized by exploring large symptom-pools for underlying constructs, without requiring that the studied subjects meet particular DSM-IV diagnoses. This line of research focuses on the constructs of anxiety and depression and their overlap (e.g. Nitschke, Heller, Imig J.C., McDonald P., & Miller, 2001). The goal of these studies is to find dimensions that assess the relevant and distinguishable aspects of the psychopathology of mental disorders, commonly seen in outpatient care. In this line of research the DSM categories do not play a role in validating the constructs, and therefore the main advantage of this approach is that it can circumvent or bypass the problems inherent to the DSM-IV.

With psychometric techniques such as exploratory and confirmatory factor analysis, the structure of latent factors that underlie the symptoms can be explored. The discovered factors could be regarded as symptom dimensions. With respect to depression and anxiety disorders (the focus of this thesis), several dimensional models have been developed. An important contribution was made by Clark and Watson with the introduction of the tripartite model (Clark & Watson, 1991). In the next paragraph (1.3) the tripartite model will be described and discussed.

1.3 The tripartite model

Description of the tripartite model

Clark and Watson introduced the tripartite model to assess and model distinct and overlapping features of mood and anxiety disorders (Clark & Watson, 1991). This symptom-based model was initially proposed to explain the high comorbidity rates of depressive and anxious disorders. The first dimension is negative affect (NA). NA is characterized by aversive emotional states, such as being distressed, fearful and nervous, and is regarded as common to both mood- and anxiety disorders. The second is positive affect (PA) and contains 'enthusiasm, excitement and energy'. A low score on this dimension is typical for depression. The third dimension is 'somatic arousal' (SA) and is specific for anxiety.

The Mood and Anxiety Symptom Questionnaire MASQ

Clark and Watson designed a 90-item instrument to measure the three dimensions of the tripartite model, the Mood and Anxiety Symptom Questionnaire MASQ (Watson & Clark, 1991). Although the model consists of three dimensions (PA, NA and SA), the items of the MASQ are allocated on five subscales; in the proposed structure of the MASQ, Watson and Clark divided the dimension NA of the tripartite model into three subcategories: General Distress Depression, General Distress Anxiety and General Distress Mixed (see Figure 1.1). The items were assigned in line with their similarity with DSM-IIIR criteria for mood or anxiety disorders (American Psychiatric Association, 1987). All 90 items are presented with a five-point Likert scale ranging from 1 (not at all) to 5 (very much).

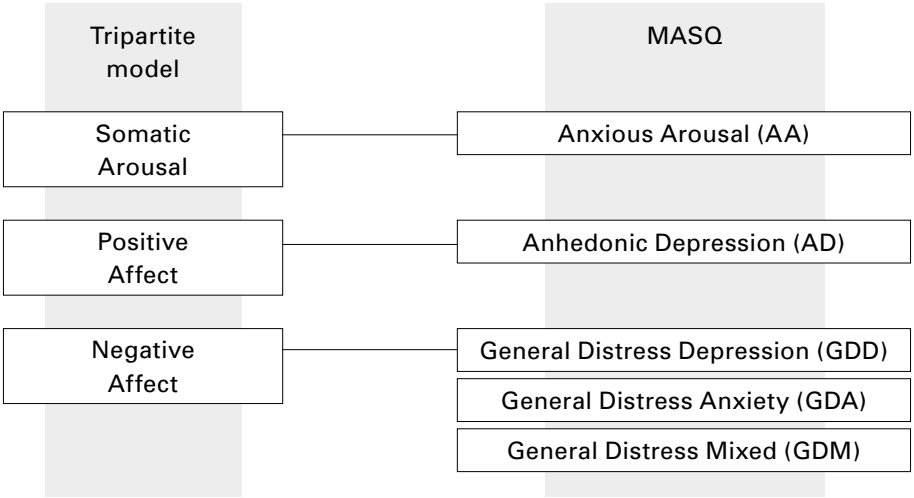


Figure 1.1 The tripartite model and the Mood and Anxiety Symptom Questionnaire.

Critique on the tripartite model

The tripartite model has inspired a large body of research, but has met some criticism as well. We will review the major points of critique on each of the three factors of the tripartite model and the MASQ.

First, there is criticism on the general distress factor NA. According to the theory of the tripartite model one would expect comparable relationships between NA and each of the different mood and anxiety disorders. However, large differences are found (Brown, Chorpita, & Barlow, 1998). This may be explained by the heterogeneous nature of NA. In the tripartite model, NA is characterized by aversive emotional states, such as feeling upset, angry, guilty, afraid, sad, scornful, and disgusted. Because the factor NA contains all these different constructs, it may be too heterogeneous to be regarded as a unidimensional construct (Shankman & Klein, 2003). With the proposed structure of the MASQ, Watson and Clark already acknowledged the heterogeneity of NA, as they divided the general distress items into three subcategories (See Figure 1.1). These three subscales of NA, however, were not confirmed with factor analysis of the MASQ-data (Boschen & Oei, 2006; Buckby, Cotton, Cosgrave, Killackey, & Yung, 2008; Burns & Eidelson, 1998).

Second, the PA scale is presented as a dimension with two endpoints: a low positive affect endpoint (lack of interest) and a high positive affect endpoint (feeling good). This assumption of the tripartite model that PA comprises of a single dimension with two extremes (lack of interest and feeling good) has however not been confirmed in analyses. Several studies showed that these two elements appear as separate factors in factor-analysis, suggesting that the two constructs do not belong to a single dimension (Keogh & Reidy, 2000; Nitschke et al., 2001; et al., 1995). Other evidence that lack of interest and feeling good are separate constructs is found in the research of Tomarken (2004). He reported medication to have a differential effect on lack of interest and on feeling good. Items assessing the low positive affect pole of the anhedonia dimension were more sensitive to earlier/lower dose bupropion SR treatment, whereas items assessing the high positive affect pole were more sensitive to later/higher dose bupropion SR treatment.

Finally, there is criticism on the third scale: somatic arousal. Although the model suggests a high score on SA is characteristic for all anxiety disorders, this is not the case, since it predominantly entails the symptoms of a panic disorder, such as increased heartbeat, transpiration, respiratory en gastro-intestinal symptoms (Chorpita & Daleiden, 2002; Joiner et al., 1999; Mineka, Watson, & Clark, 1998). Other types of anxiety symptoms like fear and worrying are therewith not well represented in the tripartite model.

The critics described above are largely focussing on the conceptualization of the tripartite model by means of the MASQ, and not on the model itself. The tripartite structure of the model is confirmed in many studies in different

populations (Bedford, 1997; Reidy & Keogh, 1997; Watson et al., 1995), and PA and NA are broadly accepted as a useful representation of affect, and being the dominant dimensions of self-reported mood (e.g., Watson, Wiese, Vaidya, & Tellegen, 1999; Remmington, Fabrigar, & Visser, 2000; Russell & Carroll, 1999). Therefore, despite the critics, the tripartite model is a strong basis for future research.

1.4 Extending the tripartite model

The current challenge is to develop a more refined model of phenotypes of psychopathology of depression and anxiety disorders, than the tripartite model. Although the tripartite model has met much support, we mentioned that it does not cover all relevant anxiety symptoms and that its negative affect dimension is rather unspecific. It is worthwhile trying to extend the tripartite model and move away from the DSM-IV. This is in concordance with the suggestions made by the original developers of the tripartite model. They proposed to specify the nature of unique components of mood- and anxiety disorders more precisely in a future model. In addition, they suggest to *"view individual disorders as representing unique combinations of different types of symptoms, with each type showing varying degrees of nonspecificity and with no type being entirely unique to any single disorder"* (Mineka et al., 1998).

Subsequently, we believe that in an ideal integrative approach, each dimension should cover a separate construct, and together the dimensions should assess the main aspects of the psychopathology of both depression and the specific anxiety disorders. The constructs should be unidimensional, have low intercorrelations, and together provide each patient with an accurate symptom profile. For instance, the two patients introduced in Textbox 1.2 would each have a different profile, despite the fact that both meet the criteria of the same DSM-diagnosis. With a specific symptom profile for each patient, more justice would be done to the idiosyncrasies of individual patients and their symptoms. Moreover, a dimensional model has the potential to associate psychopathology more accurately to etiological factors and biological markers. The ultimate goal is to improve our understanding of common mental disorders. A better understanding consequently may enable us to devise better treatments and ultimately may make it possible to predict which patients will benefit from a specific treatment for their complaints.

In this dissertation we take a step forward from the use of categorical diagnostic systems towards a dimensional approach of the psychopathology of depression and anxiety disorders. Towards this goal, several steps have to be taken. A first condition was the availability of large data sets from representative clinical samples to perform the analyses on. Until now, much of the research on dimensions of psychopathology has been undertaken with inappropriate samples such as college students or homogeneous patient groups, selected

to partake in clinical trials. Data on real life patients groups are much needed. In the spring of 2002, the Department of Psychiatry of the Leiden University Medical Center (LUMC) and the mental health institute 'Rivierduinen' (together serving a region of more than 1 million people), started collaboration for routinely assessing diagnosis and complaints at intake, and at 3-4 months intervals during treatment. The method is called Routine Outcome Monitoring (ROM). The data collected through ROM were available for this project.

The second step was to construct a Dutch translation of the MASQ, and to investigate its psychometric qualities and the fit of the tripartite model in our Dutch patient samples.

The third step was to expand the tripartite model in order to overcome some of the disadvantages of the tripartite model. Specifically, to cover anxiety more adequately additional anxiety-like constructs (e.g. phobic fear and anxious apprehension) had to be incorporated in the symptom-based dimensional model. On top of that, the statistical unidimensionality of factors had to be determined to guarantee that there were no heterogeneous dimensions in the model. We tested the applicability of various extended multidimensional models in a large set of data of psychiatric patients.

1.5 Outline of the studies

The data used for the analyses in this thesis, are collected in daily clinical practice by Routine Outcome Monitoring (ROM). Chapter 2 describes ROM as it was implemented in Rivierduinen and the LUMC in Leiden and presents data of the patient group included.

Chapter 3 describes a study into the relationship between comorbidity and severity. The aim of the study was to determine in a large outpatient sample whether patients with comorbidity have increased symptom severity and greater functional impairment as compared to patients with only a depressive or an anxiety disorder.

Chapter 4 reports on the psychometric properties of the Dutch adaptation of the Mood and Anxiety Symptom Questionnaire (MASQ). The Dutch version was constructed, its reliability and validity were tested, as was the fit of the tripartite model in our Dutch patient sample.

In Chapter 5 a proposal for an extension of the tripartite model is presented. The aim of this study was to develop scales that assess symptoms of depression and anxiety, that can adequately differentiate between depression and anxiety disorders, and that also better account for the diverse phenomenology of the various anxiety disorders.

With the study described in Chapter 6, we aimed to integrate aspects of several models into one broad dimensional model, that does not take DSM-IV diagnoses as a point of departure. The intended result was a multidimensional model to characterize patients in terms of their specific symptom profile, which

has potential for both clinical diagnosis and for use as a clinical phenotype.

Chapter 7 summarizes the main findings, provides a general discussion of these findings, and presents implications for theory, future research, and clinical practice.

CHAPTER 2

ROUTINE OUTCOME MONITORING IN THE NETHERLANDS: PRACTICAL EXPERIENCES WITH A WEB-BASED STRATEGY FOR THE ASSESSMENT OF TREATMENT OUTCOME IN CLINICAL PRACTICE

Clinical Psychology and Psychotherapy 2011; 18, 1-12.

Edwin de Beurs
Margien E. den Hollander-Gijsman
Yanda R. van Rood
Nic J.A. Van der Wee
Erik J. Giltay
Martijn S. van Noorden
Rosalind van der Lem
Esther van Fenema
Frans G. Zitman

Abstract
Routine Outcome Monitoring (ROM) is a method devised to systematically collect data on the effectiveness of treatments in everyday clinical practice. ROM involves documenting the outcome of treatments through repeated assessments. Assistants are employed who perform a baseline assessment comprising a standardized diagnostic interview, administration of rating scales, and completion of several self-report measures by the patient. At fixed time intervals assessments are repeated. Dedicated web-based software has been developed to assist in this task. ROM informs therapists and patients on the severity of the complaints at intake, and the waxing and waning of symptoms over the course of treatment. Researchers can use ROM for effectiveness research and managers can use it for benchmarking. The use of ROM for research is illustrated by presenting data on the diagnostic status of patients participating in ROM and data on treatment outcome data of a subgroup of patients (with panic disorder) in our database. The results show that implementation of ROM is feasible and, after some initial reservations, most therapists now consider ROM to be a necessary and important adjunct to the clinical treatment. In addition, ROM furthers research as the data can be used to study the phenomenology of psychiatric disorders and the outcome of treatments delivered in everyday practice.

Key Practioner Message:
- A form of tracking the progress of treatment through Routine Outcome Monitoring (ROM) is described.
- Implementation of ROM appears feasible and can be carried out in large institutions as well as smaller practices.
- Providing feedback about outcome in an appealing format is highly valued by both therapists and patients.
- ROM data enable investigation of the effectiveness of treatments in everyday clinical practice.

2.1 Introduction

Routine Outcome Monitoring (ROM) is the assessment of treatment outcome at regular intervals in order to monitor patients' progress during treatment. It involves the application in everyday clinical practice of assessment technology that was originally developed for randomized clinical trials. Several objectives may be achieved with ROM. It provides information on type and severity of psychopathology before treatment commences, which can be used to optimize allocation of patients to treatment forms. Further, ROM provides feedback to therapist and patients on progress made in treatment. Finally, ROM data can be used for research into the effectiveness of treatments in care as usual.

Already in 1988, Ellwood proposed routine and frequent assessment of patients' health and suggested to build large databases from these data (Ellwood, 1988). Although this idea was well-received in editorials (see, for recent examples, Holloway, 2002; Slade, 2002), in clinical practice routine assessment is seldom realized. In a survey among 396 psychiatrists in England, only 19.4% *"routinely or occasionally"* measured the outcome of the treatment provided (Gilbody, House, & Sheldon, 2002). Since then, some projects have been initiated in which treatment outcome is routinely assessed using different outcome measures.

In the UK two developments are worth mentioning, the Mental Health Minimum Data Set (MHMDS) of the National Health Service (http://www.ic.nhs.uk/services/mental-health/mental-health-minimum-dataset-mhmds) and the Clinical Outcomes in Routine Evaluation (CORE) system. Since 2003, mental health institutes are required by the Department of Health to provide anonymized outcome data on treatments to the MHMDS. The HoNOS is the central part of the MHMDS. The HoNOS was developed as a clinician-rated instrument for routine outcome assessment and appeared a reliable, valid and sensitive outcome measure, especially suitable for the more severe mental disorders (Wing et al., 1998). Until now, the NHS has reported only results on data quality and no outcomes on "spells of care" have been reported yet. The CORE system was designed as a quality evaluation system to evaluate therapy service delivery. Its central measure, the CORE-OM, was developed as a *"user-friendly, pantheoretical and free measure to monitor the outcome of counseling and psychotherapy"* (Barkham, Culverwell, Spindler, & Twigg, 2005). It is best suited for the less severe, more common mental disorders, such as mood and anxiety disorders. Stiles et al. (2006) report on its application in evaluating the outcome of various treatments of patients that were mostly seen in primary care. Interestingly, they found a large treatment effect (average effect size [ES]=1.36 for the pre-post difference), but little difference in outcome was found between theoretically different approaches to treatment.

In Australia a nationwide program of routine outcome measurement was implemented in 2000 (Burgess et al., 2009). Mental health services are

required to provide outcome data for a national database. In this program both clinician-rated (HoNOS) and self-report instruments (e.g., K-10+,Kessler et al., 2003) are used. To analyse the data, the Australian Mental Health Outcomes and Classification Network (AMHOCN) was established in 2003. They not only analyse and report individual and aggregated results (benchmarking) but also take steps to organize the data properly, and give trainings on how to collect and use data.

In the USA, Lambert, Hansen and Finch (2001) coined the term *"patient-focused"* research for routine assessment of the course of symptoms over time. They promote session-by-session assessments and developed a relatively short questionnaire for this purpose: the Outcome Questionnaire (OQ-45, Lambert et al., 1996). The results of the OQ-45 are discussed with the patient, allowing ample time for this in the session. The high frequency of assessments makes short-term changes in psychopathology and functioning visible, but limits the number of items that can be administered and thus the comprehensiveness of the assessment. Apart from reporting on the changes per session, the expected trajectory of scores at future assessments is estimated using the pretreatment score. If a patient's score falls outside a specified range around the projected score the therapist receives a warning of potential premature dropout and/or negative outcome should therapy continue unchanged. A similar approach is advocated by Miller and colleagues (Miller, Duncan, Sorrell, & Brown, 2005). They use an even shorter scale, the Outcome Rating Scale (ORS), comprising only four visual analog scale items to be completed every session. These four items mirror the four subscales of the OQ-45. In addition, at the end of the session the patient also rates the therapeutic alliance and the usefulness of the session (agreement on goal, methods, and overall approach of therapy) on a Session Rating Scale comprising also four visual analog scales. This score is used to assess whether discrepancies exist between what a patient wants from therapy and is receiving (Miller et al., 2005).

This paper presents a method for monitoring treatment outcome in clinical practice which has been implemented in the Netherlands. In contrast to projects described above, we employ a less frequent but more comprehensive assessment battery including both generic and disorder-specific measures, and evaluate treatment outcome from the viewpoint of the patient and an independent rater. The method of ROM is described, as are the experiences with ROM and the use of ROM data by managers and researchers. To illustrate the potential of ROM for research purposes we present the baseline characteristics (diagnoses and comorbidity) of the first cohort of patients, and present outcome data of a subgroup of patients with panic disorder.

2.2 Method

General description

In spring 2002, the mental health clinics of 'Rivierduinen' (an institute serving a region of more than 1 million people) and the Department of Psychiatry of the Leiden University Medical Center (LUMC) started collaboration for routine assessment of the diagnosis at intake, and the severity of complaints at intake. Reassessments take place every 3-4 months during treatment. ROM is restricted to patients referred for treatment of mood, anxiety, and somatoform (MAS) disorders. These patients form a relatively homogenous group with substantial mutual comorbidity (Kessler et al., 1996) and mainly receive outpatient care. So far, patients referred for treatment of other disorders, such as addiction or substance abuse or Axis II disorders, do not participate in ROM. Finally, to be included patients must be literate and have sufficient command of the Dutch language to complete the self-report instruments. For the present report, a group of patients with complete data was selected (N=3,798) and their data were analysed to illustrate the research potential of ROM- data.

Ethical considerations and privacy issues

At intake, patients are informed that ROM is part of the general policy of Rivierduinen to monitor treatment outcome, that outcomes are made available only to their therapist, and that the data will be used for research purposes, but only in anonymized form. If patients object to such use, their data are removed. A comprehensive protocol safeguards anonymity of the patients and ensures proper handling of the data. This protocol is available for patients on request. The Medical Ethical committee of the LUMC approved the regulations and agreed with this policy.

Instruments

At intake the Axis-I diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) is established using the Mini-International Neuropsychiatric Interview-plus (MINI-plus, Sheehan et al., 1998). The Dimensional Assessment of Personality Pathology (DAPP-SF) is administered to assess maladaptive personality traits (Livesley & Jackson, 2006; Van Kampen D., De Beurs E., & Andrea, 2008). Subsequently, a number of instruments are administered at intake, which are also completed at each re-assessment to allow for the evaluation of treatment outcome. Together, these instruments cover change in three areas of functioning: symptom reduction, increased wellbeing, and improvement in general life functioning (Sperry, Brill, Howard, & Grissom, 1996). They are commonly used in treatment outcome research and have good psychometric properties as evidenced by national and international publications (an overview of instruments used is available at http://www.lumc.nl/psychiatry/ROM-instruments). Outcome is assessed

by patients' self-report and by an independent assessor, and includes both generic and disorder-specific measures. Generic instruments are completed by all patients, e.g., the Brief Symptom Inventory (BSI, Derogatis, 1975) and allow the comparison of treatment outcome among all patients irrespective of their disorder. Disorder-specific measures are administered only to those patients meeting criteria for the disorder at hand, e.g., the Beck Depression Inventory Revised (BDI-II, Beck & Steer, 1987) in case of a mood disorder. The latter instruments are more sensitive to change as they assess the intensity of the symptoms which the treatment targets; they provide a more accurate picture of the clinically important improvements or progress of the individual patient (Lee, Jones, Goodman, & Heyman, 2005). The assessment outcomes are made available to the therapist, discussed with the patient, and used to support decision-making for the future course of the treatment.

Specialized staff for ROM

The LUMC and Rivierduinen employ and train psychiatric nurses and psychologists (Master's level) to carry out ROM. They are less costly than therapists, and ratings from a small specialized staff tend to be more reliable than ratings from therapists. ROM assistants administer the MINI-Plus interview, rating scales such as the Clinical Global Impression (CGI, Guy, 1976) and the Global Assessment of Functioning (GAF, Endicott, Spitzer, Fleiss, & Cohen, 1976), and write a brief report (1-2 pages) on the main findings for the therapist.

To date, 20 ROM assistants have been trained in the administration of the MINI-Plus interview and the rating scales. Initially, weekly training sessions were organized. From 2006 on, assistants who had started in 2002 and had at that time ≥ 4 years experience with the ROM instruments, assisted in training new staff. To further sharpen their diagnostic skills, ROM assistants currently still meet every month (for half a day) for instruction (by invited speakers) on the phenomenology of various disorders. In addition, they practice with rating scales to improve interrater reliability. Videotaped interviews with patients are rated and the ratings are afterwards compared and discussed to reach consensus.

Treatments

The diagnostic information from the first ROM assessment is used in conjunction with the standard clinical intake interview to select the optimal treatment for the patient. Psychiatrists and clinical psychologists provide treatment in accordance with the national multidisciplinary guidelines of the National Steering Committee describing evidenced-based treatments for mood and anxiety disorders. Treatment usually consists of medication, mainly selective serotonin reuptake inhibitors (SSRIs), cognitive behaviour therapy (CBT), or

a combination of both. Simultaneous with the start of ROM, a new stepped-care approach to treatment delivery was introduced in which the first treatment of choice is the least invasive/least intensive treatment for which efficacy has been established (e.g., a protocolled CBT or short course of pharmacotherapy). Only when this treatment does not result in sufficient symptom reduction, a more invasive or intensive treatment is offered (e.g., a combination of CBT and pharmacological treatment or, eventually, electroconvulsive therapy).

Clinically significant change

To designate a change from pretest to retest as clinically meaningful we follow the proposal of Jacobson and colleagues (Jacobson, Roberts, Berns, & McGlinchey, 1999) to combine two statistical criteria for clinical significant change. First, the change from baseline to posttest should fall outside the range of the measurement error of the instrument, i.e., a statistically reliable change should be attained (Reliable Change Index). Secondly, the posttest score should be beyond a cut-off point signifying the transgression from dysfunctional to functional, i.e., a clinically significant change. Combining these two criteria provides five possible outcomes: recovery (both criteria are met), mere improvement (only statistically reliable change), no change, deterioration (reliable change in the 'wrong' direction), and relapse (reliable change and a posttest score which falls within the dysfunctional range). In ROM the results of all instruments and subscales are provided in terms of these five possible outcomes.

Feedback to the therapist and patient

ROM provides the therapist with detailed information on the state and progress of their patients. The therapist shares and discusses these results with the patient. The report on the baseline assessment consists of a summary of the results of the diagnostic interview and a selection of the most relevant results of the instruments (Figure 2.1). The re-assessment report describes the progress made since the previous assessment, or presents a review of the course of complaints over successive assessments (Figure 2.2). To accommodate therapists and patients, care is taken to present results in a visually attractive way and to provide feedback without delay. Reports follow within one day of the (re)assessment.

Therapists use the reports to inform their patients about the results. Patients are not granted direct access to their data; it was considered important to assist and inform patients on how to interpret the results in an appropriate way. The results (such as depicted in Figs. 2.1 and 2.2) can be printed and given to the patient to take home. Apart from being used to inform patients, the ROM results are also used in staff meetings were the course of treatment of patients is discussed periodically.

Brief Symptom Inventory (BSI), Assessment 1, (12-12-2008)

Personal data:

Name:	J. Doe
Birth year:	1969
Reg. number:	0321431
Therapist:	E. de Beurs

Results:

			compared to:	
subscale	score	miss	patients	gen. population
SOM	61 #	0%	high	very high
OC	50 #	0%	average	high
I S	46 #	0%	below av.	above av.
DEP	38	0%	low	average
ANX	60 #	0%	above av.	very high
HOS	49 #	0%	average	high
PHO	65 #	0%	high	very high
PAR	48 #	0%	below av.	above av.
PSY	57 #	0%	above av.	very high
TOT	52#	0%	above av.	very high

the score is higher than the cut-off between functional/dysfunctional

In a graph:

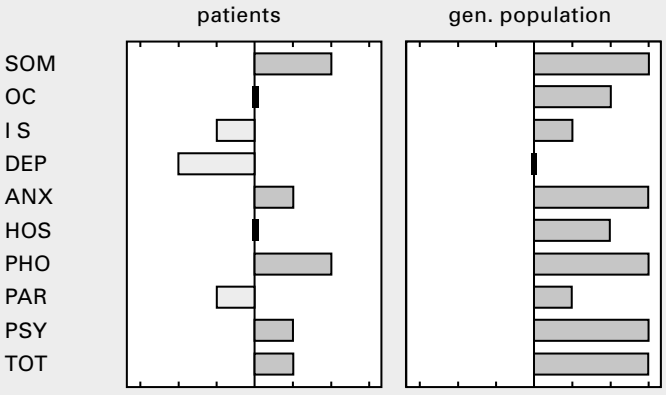


Figure 2.1 An example of the output from a single assessment with the Brief Symptom Inventory (T-scores).

Brief Symptom Inventory (BSI) - course

Personal data:

Name:	J. Doe
Birth year:	1969
Reg. number:	0321431
Therapist:	E. de Beurs

Results:

	Assessment:						
subscale	1 (12-12)		2 (4-9)		3 (8-20)		4(11-22)
SOM	61	>>	42	-	42	-	46 (below av.)*
OC	50	-	39	-	38	-	39 (low)
I S	46	-	44	-	37	-	37 (very low)
DEP	38	-	35	-	35	-	35 (very low)
ANX	60	>	44	-	41	-	39 (low)
HOS	49	-	42	-	42	-	42 (low)
PHO	65	>	48	-	50	-	42 (low)
PAR	48	-	43	-	41	-	39 (very low)
PSY	57	>>	37	-	40	-	40 (low)
TOT	52	>>	39	-	37	-	37 (low)

* normative level of the last score
- no significant change
> improved (sign. change in comparison to the previous assessment)
>> recovered (significant progress and transgression of the cut-off)

In a graph:

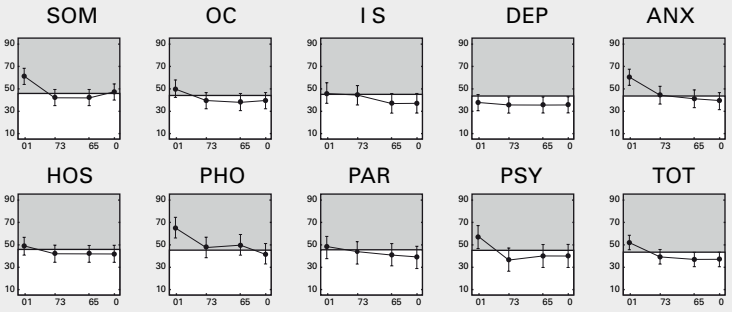


Figure 2.2 The course of complaints according to the Brief Symptom Inventory (T-scores) over time.

Software support

Dedicated web-based computer software has been developed for the administration of the MINI-Plus diagnostic interview, completion of rating scales, administration of self-report measures, and ascertainment of treatment outcome. The software presents each question of the MINI-Plus on the screen of the interviewer together with the response options. The computer software is able to deal with the sometimes complicated scoring rules in this interview and is 'intelligent': if sufficient symptoms are answered as absent to preclude a diagnosis, or sufficient symptoms are rated present to establish a positive diagnosis, no additional questions are asked, after which the module is closed and the next module is started.

The software is also used for completion of self-report questionnaires. For this purpose it has been designed as an open system: any questionnaire can be defined and administered with the software. The assessment can take place at the clinic where touchscreens can be used to accommodate computer-illiterate patients or, if they wish, patients can complete questionnaires at home via the internet.

The software computes (sub)scale scores and compares them with normative values for male/female *patients* and male/female respondents from the *general population* (Figure 2.1) and depicts the course of symptoms over time (Figure 2.2). Furthermore, the software helps in the management of data collection, e.g. allowing to list all patients who need to be invited for an upcoming 'outcome assessment' session. Finally, the software allows for the export of aggregated and anonymized data for analysis with statistical software, such as the Statistical Package for the Social Sciences (SPSS).

2.3 Results

Experiences with ROM

Interrater reliability

Multiple assessments of the same case were available from the training sessions of the ROM staff, but these cannot be used to formally establish interrater reliability, as this would lead to underestimation of the reliability of experienced ROM assistants. Interrater reliability has, however, been formerly assessed with a small subset of patients (n=44) revealing sufficient interrater reliability for the Comprehensive Psychiatric Rating Scale (Goekoop, De Beurs, & Zitman, 2007); average Cohen's κ=0.60), the GAF (average Cohen's κ=0.73) and the CGI (average Cohen's κ=0.55). These indices denote acceptable interrator reliability.

Time investment

The time needed for the first assessment is about 2 hours; 35 minutes for the MINI-Plus, 40 minutes for the rating scales and 45 minutes for the self-

report measures. A ROM re-assessment session takes (on average) 1 hour to complete. Research assistants, however, reported that for some patients there was insufficient time to include all the disorder-specific instruments which should be administered according to the MINI-Plus diagnoses.

Acceptability of the ROM procedure to patients

Patients showed good compliance with the ROM procedure. The percentage of patients with a mood, anxiety, and/or somatoform disorder that participated in ROM increased to 80% by 2009. Reasons for not participating were: the patient's command of the Dutch language was deemed insufficient to complete the questionnaires, or the assessment procedure was considered too invasive for the patient. No patients refused to partake in the ROM procedure, but approximately 5% failed repeatedly to show up at their first assessment. Comparison of the demographic data of patients who did and did not participate in ROM revealed no significant difference for gender, age, or educational level (all $p > 0.20$); however, more patients with a non-Dutch ethnic origin did not participate in ROM.

Patients were satisfied with ROM; they did not feel excessively burdened and the comprehensive assessment made them feel that their problem or disorder was taken seriously by the staff.

Attrition

Even though patients were willing to participate in ROM at intake, in our study sample, on each successive assessment the cohort was reduced in size by about 50%. Half of the patients without a re-assessment had discontinued their treatment and their last assessment can be considered a proper endpoint. However, the other half was still undergoing treatment, should have been assessed, and is considered as real loss to follow-up. Thus, no formal endpoint assessment was available for about 25% of the patients of the baseline sample due to repeated no-show. The assessment session had been rescheduled twice for these latter patients before we gave up on their outcome data. No-show for re-assessment ranges from 10 to 30% of the appointments, making it a costly problem.

Therapists' impressions of ROM

In an early phase of the implementation of ROM we conducted a survey among therapists and managers, investigating their views on the accuracy and usefulness of the data in their day to day clinical work. Therapists reported that they utilized the outcome data to motivate patients by showing them the progress made thus far, and the symptoms that still need attention in treatment. Initially, some resistance from therapists toward standardized assessment had to be overcome. Some felt that ROM was intrusive, violating the privacy of the

therapy dyad. Others felt they were better able to judge the clinical progress than could be done with standardized instruments. In practice, however, it appeared that ROM data supported or supplemented their clinical impression on how the patient fared in treatment; the data sometimes even corrected a false impression. As a result, therapeutic staff became more sensitive to treatment outcome data and eventually the majority enthusiastically accepted ROM.

During staff meetings the ROM results are presented and when they demonstrate lack of progress different courses of action are discussed. Likewise, when the ROM results indicate recovery, i.e., reliable and clinical relevant decreases in scores measuring the intensity of the main complaints, this signals that therapy might be ended, preventing the unnecessarily lingering on of treatment. Thus, ROM more than likely improved the efficiency of the treatments provided in the clinic.

Other use of ROM data

Apart from therapists and patients, researchers and managers may also use ROM data. Managers have just started to use ROM data for internal benchmarking purposes. As yet, only results on the proportion of successfully monitored treatments have been compared among the seven outpatient clinics. Outcomes on differential effectiveness of various treatment programs, locations, departments, or even therapists, have so far not been supplied. For these outcomes more complete data are needed, i.e. less loss of re-assessments.

ROM data have been used for psychometric research (Wardenaar et al., 2010; De Beurs E., Rinne, van, Verheul, & Andrea, 2009; De Beurs, Den Hollander-Gijsman, Helmich, & Zitman, 2007; Den Hollander-Gijsman, De Beurs, Van der Wee, Van Rood, & Zitman, 2010; Van Kampen D. et al., 2008), treatment outcome research, and for basic research (Van Noorden et al., 2010; Veen et al., 2009).

Examples of findings with ROM data

The results described in this paper are based on ROM data collected from January 2004 to December 2006. This dataset consists of 3,798 patients. The average age of the group was 39.6 (SD=13.3) years and 63% were women.

Diagnostic status at intake

According to the MINI-Plus, 1,618 patients (42.6%) met criteria for one MAS disorder, and 1,556 patients (41.0%) had more than one concurrent disorder (967 patients (25.5%) with two comorbid disorders, 403 patients (10.6%) with three, and 186 patients (4.9%) with four or more). Figure 2.3 presents an overview of the various (combinations of) diagnoses found in this sample when grouped in higher-order categories of MAS disorders: 1,788 patients (47.0%) met criteria for one or more *mood* disorders, 1,653 patients (43.5%) for one or more *anxiety*

disorders, 653 patients (17.2%) for one or more *somatoform* disorders, and 851 patients (22.4%) had other disorders (e.g., adjustment disorder, mixed anxiety-depression), or did not meet criteria for a DSM-IV Axis-I diagnosis (16.4%).

Outcome for panic disorder patients

To further illustrate the potential of ROM, we investigated the treatment outcome of patients with panic disorder using the Panic Disorder Severity Scale (PDSS) as an observer rater instrument for the assessment of the intensity of panic disorder symptoms (Shear et al., 1997). A total of 415 patients had a MINI-Plus diagnosis of panic disorder (with current panic attacks) and filled in the PDSS. Their average age was 35.9 (SD=10.7) years; 64% were women and 62% suffered from panic disorder with agoraphobia. On the PDSS the average score at pretest was 12.34 (SD=5.03). A second assessment after (on average) 25.8 (SD=18.7) weeks was available for 238 patients. In this subsample, the PDSS total score dropped from 12.56 (SD=5.01) to 7.04 (SD=5.84), a difference of almost 1 SD. At posttest 68% scored below the cutoff score of 7 on the PDSS indicating clinically significant change (at the pretest 18% of the patients scored below 7). Finally, 58 patients completed four assessments spanning (on average) 62 weeks of treatment. Multivariate analysis of variance for repeated measures was used to test whether the drop in score over time followed a linear pattern: $F_{\text{linear contrast}}(1, 57)=52.11$, $p < .000$, partial $\eta 2=.48$, which denotes a large effect.

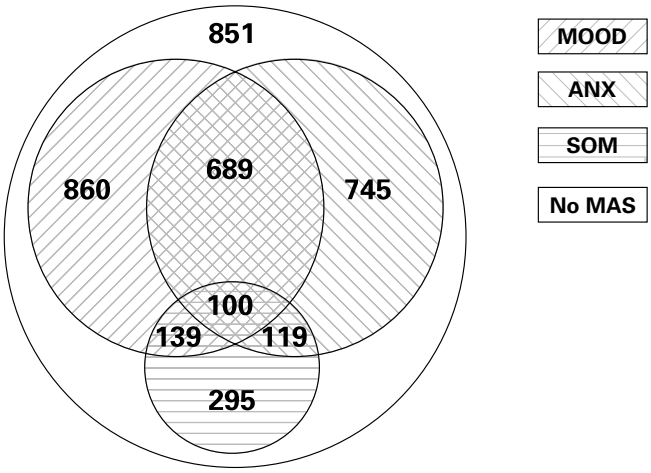


Figure 2.3 Number of patients with mood (MOOD), anxiety (ANX), somatoform (SOM) disorders (or MAS disorders), and those not meeting criteria for mood, anxiety, or somatoform disorders (No MAS) and their pattern of comorbity.

2.4 Discussion

The value of ROM

Our experiences with ROM suggest that ROM offers several benefits; however, these need to be investigated in more detail (preferably in a controlled study) as the present study provides no empirical proof of the positive value of ROM. In the literature different approaches to ROM have been developed and described. Some use a single instrument (e.g., the HoNOS, CORE-OM, or the OQ-45), others use multiple measures from different perspectives (e.g. clients, therapist or clinical raters). Some administer these instruments at pretest and posttest only, others assess periodically, and still others (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005) organize assessments on a session-by-session basis.

The benefits from monitoring and providing feedback to patients have been studied in a few studies using a controlled design. Lambert and colleagues (2005) compared three conditions: informing both the therapist and the patient of the results, informing only the therapist, and informing neither the therapist nor the patient. They report positive effects on treatment outcome of providing feedback. The greatest reduction of symptoms was seen in the condition where both parties were provided with feedback. In this condition the rate of patients demonstrating clinically significant improvement was doubled. Slade et al. (2006) evaluated in their controlled study the effects of 3-monthly feedback to patients treated in a community mental health centre. These patients regularly completed (postal) questionnaires on their mental health. On the primary outcome measures the intervention group did not fare better than the 'treatment as usual' control group. However, informed patients spent significantly less days in in-patient care which made the intervention cost effective. A recent meta-analysis on the effects of providing feedback on therapist and patients concluded that the benefits are present but rather limited in effect size (Knaup, Koesters, Schoefer, Becker, & Pushner, 2009). The research of Lambert and colleagues (Lambert et al., 2005) and Miller and colleagues (Miller et al., 2005) suggests the best results are attained with patients that otherwise would have stopped the treatment prematurely. Thus, additional controlled studies are needed, in view of the substantial efforts and costs involved in obtaining outcome data in a routine manner. In addition, further research is required to determine the minimal assessment battery necessary to serve all stake holders; i.e. therapists and patients, researchers and managers.

Attrition

In our study sample, on each successive assessment the cohort was reduced in size by about 50%. High attrition rates with ROM are frequently reported. For example, in the study of Stiles et al. (2006) posttest data were available for only 33% of the patients that had provided pretest data. The high number of

patients that is lost for follow up precludes conclusions on the effectiveness of the treatments evaluated (Clark, Fairburn, & Wessely, 2008). To increase the number of patients with complete ROM data we have improved communication between ROM assistants and therapist. Now, therapists are required to inform the ROM assistant if treatment is about to conclude. The assessment session can then be scheduled prior to the final treatment session. Until more complete data are available, an intention-to-treat analysis of aggregated data might yield a more valid reflection of results obtained in everyday clinical practice than is provided by a completer analysis (Wood, White, & Thompson, 2004).

Ratings versus self-report data

ROM data are (in part) based on ratings by the ROM staff. Reliability of these ratings ranged from 0.60 to 0.74 (Cohen's κ). This denotes moderate to substantial agreement between the raters. However, these interrater reliability estimates may be somewhat inflated as they are based on the re-rating of a videotaped interview. Subjecting patients twice to two separate interviews was deemed too demanding for patients, but would have yielded more valid (and likely lower) interrater reliability estimates.

Traditionally, outcome research with mood disorders relies predominantly on rating scales whereas with anxiety disorders it is more common to use self-report scales. Using the same instruments in ROM allows for direct comparison between the treatment results attained in clinical practice (efficacy) and in randomized controlled studies. However, there are considerable additional efforts and costs involved in utilizing ratings made by independent observers. The incremental value of using raters and rating scales, compared to assessing outcome by patients' self-report with questionnaires needs to be further investigated.

Computerised assessment

We decided to use computerized administration of questionnaires for ROM as this implies that there are no missing data, instruments are scored straight away, and normed results are immediately available. A disadvantage is that some experience with computers is required, which older people in particular might not have. To solve this problem the assessment sessions are scheduled at the treatment centres were touchscreens are available and ROM staff can help the patient when this is necessary.

The software allows for completion of the self-report questionnaires at home. Although this is patient-friendly, the drawback is that we cannot be 100% certain that the patient completed the measures without the help of family members or others. The option of completing self-report questionnaires at home is, however, still open and we are currently exploring options for this more patient-friendly version of ROM.

Scientific research implications

To illustrate the research potential of ROM data, we investigated the diagnostic data of all the patients participating in ROM between January 2004 and December 2006, and the treatment outcome of panic disorder patients. The diagnostic data of the MINI-Plus reveal that in clinical practice comorbidity abounds: a large proportion of patients meet criteria for two or more diagnoses. This may in part be due to the use of a structured diagnostic interview in which the criteria of a large number of DSM disorders are methodically checked and diagnoses are not easily overlooked. In a clinical interview, the intaker might be more focused on establishing the disorder to treat, disregarding comorbid psychopathology, and may see symptoms which could qualify for a comorbid diagnosis as belonging to the primary diagnosis (See also Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009). The importance of diagnostic accuracy was underlined by Jensen-Doss and Weisz (2008) who showed in a meta-analysis that less drop-outs occurred and a better outcome was attained in cases where clinicians and researchers (using a structured diagnostic instrument) agreed about the diagnosis.

The patients monitored in this study form a representative sample of the patients typically seen in clinical practice. The ROM data from this sample can be used to investigate whether these patients differ substantially from patients that participate in clinical trials. Treatment outcome of panic disorder patients was assessed with a widely-used disorder-specific outcome measure, the PDSS. After (on average) 6 months of treatment scores on this rating scale had dropped by almost 1 SD. Barlow, Gorman, Shear, and Woods (2000) in their landmark randomised controlled trial, report a response rate of 60%, defined as a score below the threshold of 7 for clinically relevant complaints. In our sample, 68% of the patients scored below 7 at post-test, comparing favourably with the results of Barlow and colleagues.

The findings on comorbidity and panic disorder treatment illustrate that data collected through ROM can be used for research. Researchers normally do not have easy access to the treatment results attained in mental health institutes. With ROM the interests of both therapists and researchers are served. The ROM structure allows for the collection of additional data, such as information on biological, social, psychological, or cognitive functioning of patients. With these data fundamental research questions can be addressed regarding differences between diagnostic subgroups, associations between the phenomenology of disorders and biological or psychological parameters, and the prognostic value of these variables for treatment outcome. The latter can be advantageous for clinical practice, potentially allowing for a better match between patient needs and treatment. For instance, the choice of medication in the treatment of anxiety or depression is largely a process of 'trial and error' to find an acceptable balance between side-effects and optimal

therapeutic effect. In the future, it might be possible to select medication based on the patient's genetic information. Currently, a biological database is being built with genetic information of the patients who participate in ROM. This will enable future research into the genetic background of the phenomenology of common mental disorders and may yield preliminary data on the interaction of genes, pharmacological agents and treatment response.

2.5 Conclusions

In summary, ROM is a method for the routine assessment of treatment outcomes in clinical practice, which simultaneously serves the interests of patients, therapists, mental healthcare managers, and researchers. Implementation of ROM has been shown to be feasible and created an efficient 'assessment culture' in a mental health institute with little academic tradition.

CHAPTER 3

COMORBID DEPRESSION AND ANXIETY: SYMPTOM AND FUNCTIONAL SEVERITY IN THE CLINICAL SETTING

Submitted for publication

Margien E. den Hollander-Gijsman
Nic J.A. van der Wee
Edwin de Beurs
Klaas J. Wardenaar
Yanda R. van Rood
Martijn S. van Noorden
Erik J. Giltay
Frans G. Zitman

Abstract

Background
Comorbidity between depression and anxiety disorders is widely understood to be associated with poorer outcome, increased symptom severity and more functional impairment. However, symptom severity and functional impairment in comorbidity have not been compared directly to those in pure depression and pure anxiety disorder in general psychiatric patient samples. The aim of this study is to determine in a large outpatient sample whether patients with comorbidity have increased symptom severity and greater functional impairment as compared to patients with only a depressive or an anxiety disorder.

Method
Analyses were performed on a large sample consisting of 2278 outpatients with a depression and/or an anxiety disorder from a general psychiatric setting. We studied the relation of diagnostic status with global severity, functional severity, depression severity and anxiety severity.

Results
Symptom severity (global severity, depression severity and anxiety severity) and functional impairment were increased in the comorbid group as compared to the pure groups. Depression severity in the comorbid group was higher than in the pure depression group and anxiety severity in the comorbid group was higher than in the pure anxiety group. The latter was also the case when analyses were repeated for specific DSM-IV anxiety disorders.

Conclusions
In a large general psychiatric outpatient sample comorbidity is associated with increased depressive and anxiety severity, and increased functional impairment.

3.1 Introduction

High rates of comorbidity between anxiety disorders and depression have been reported in the general population (Kessler et al., 1996), in primary care (Roca et al., 2009) and in secondary care (Brown et al., 2001). Comorbidity between depression and anxiety disorders is widely understood to be associated with increased severity, poorer outcome and more functional impairment. Researchers have argued that comorbidity of depression and anxiety disorders even warrants a separate diagnosis (e.g. Tyrer, 2001; Silverstone & Von Studnitz E., 2003). The appreciation of the importance of comorbidity is also reflected in the goals set for the next version of the Diagnostic and Statistical Manual (DSM)-5, as this version should enable clinicians and researchers to take the presence and effects of comorbidity into account (www.dsm5.org).

However, contrary to many studies on the prevalence of comorbidity or its effects on disease outcome (Emmanuel, Simmonds, & Tyrer, 1998), studies reporting on the severity of depressive and anxiety symptoms in comorbidity usually did not include the full spectrum of symptomatology. In these studies, only depression with and without comorbid anxiety disorders (e.g. Dalrymple & Zimmerman, 2007; Fava et al., 2004), or only anxiety disorders with and without comorbid depression (Kaufman & Charney, 2000), but never the three groups together, were examined. Furthermore, these studies have only been performed in the general population or in samples from specialized psychiatric care settings, or did assess only the severity of one type of symptoms (depressive or anxiety) or the impact of comorbidity of depression and anxiety on general distress and daily functioning (Pirkola et al., 2003; Wittchen, Carter, Pfister, Montgomery, & Kessler, 2000).

It is of relevance for routine clinical practice and for the ongoing research into the nature and consequences of comorbidity, to obtain more insight in the severity of depressive and anxiety symptoms and their effects on functioning in patients with comorbidity in routine general psychiatric practice settings. In this study, we examined the severity of depressive and anxiety symptoms and the related functional impairment of patients with comorbid and pure mood and anxiety disorders in a large naturalistic routine psychiatric outpatient sample. We hypothesized that patients with comorbidity of depression and anxiety disorders would have 1) a higher global symptom severity than patients with only one disorder, 2) more severe depressive symptoms than patients with a pure depression, 3) more severe anxiety symptoms than patients with a pure anxiety disorder and 4) a higher severity of functional impairment than patients with a pure depression or a pure anxiety disorder. Because one can expect different effects for different anxiety disorders, we also tested the third hypothesis for specific anxiety disorders.

3.2 Method

Routine Outcome Monitoring

This study was conducted on data collected through Routine Outcome Monitoring (ROM, (De Beurs et al., 2011)). ROM is an ongoing monitoring system for patient care, implemented in the outpatient clinics of Rivierduinen (a large organization for the provision of mental health care in the province of Zuid-Holland, the Netherlands) and the Department of Psychiatry of the Leiden University Medical Center (LUMC). All patients referred to these clinics for treatment of a mood-, anxiety- or somatoform disorder, have assessment sessions with a psychiatric research nurse at the start, during, and at the end of the treatment (De Beurs et al., 2011). For this study, the baseline ROM-assessments were used. During these baseline assessments, a standardized diagnostic interview is administered and interviewer and self-reported ratings are completed. ROM data are primarily used for diagnosis and to inform clinicians and patients about treatment progress. The use of anonymous data of these patients for research purposes has been approved by the Medical Ethical Committee of the Leiden University Medical Center.

Sample

The initial group consisted of 3798 outpatients admitted consecutively between January 2004 and December 2006. For the present study, three diagnostic groups of patients were selected (total n=2278): (1) patients with one or more anxiety disorders and no depression (n=729), (2) patients with a depression and no anxiety disorders (n=860) and (3) patients with comorbid an anxiety disorder and a depression (n=689). The diagnosis 'depression' includes both Major Depressive Disorder (MDD) and dysthymia (respectively 62,6% and 5.4% of the total sample (n=2278)), but not bipolar disorder. The majority of the patients that were not included in this study had a single or comorbid somatoform disorder.

Instruments

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

To establish the presence of current and life-time Axis-I disorders according to the DSM-IV diagnostic criteria, The Dutch translation of the M.I.N.I.-Plus 5.0.0-R (Van Vliet, Leroy, & van Megen, 2000) was used (Sheehan et al., 1998). The M.I.N.I.-Plus is an extended version of the original M.I.N.I. Lecrubier and colleagues (Lecrubier et al., 1997) report sufficient reliability of the M.I.N.I (k=0.88-1.00; test-retest reliability= 0.76-0.93). Validity was demonstrated by sufficient concordance with the Composite International Diagnostic Interview (CIDI, WHO). Interviews were performed by extensively trained and supervised psychiatric research nurses. All diagnoses reported in this study were current at the time of assessment. In the M.I.N.I.-Plus some hierarchical exclusion rules

apply: in case of a current depression diagnosis, concurrent dysthymia is ruled out. Depression, and generalized anxiety disorder (GAD) can only be diagnosed concurrently if both disorders have a different time of onset.

Brief Symptom Inventory
The Brief Symptom Inventory (BSI) (Derogatis & Melisaratos, 1983; De Beurs, 2005) is a shortened version of the Symptom Checklist (SCL-90) (Derogatis, Lipman, & Covi, 1973) , and is used to measure psychological complaints or symptoms. The BSI consists of 53 items, rated on a five-point Likert scale, ranging from 0 (not at all) to 4 (very much). The items are subdivided into nine subdimensions: 1) somatic complaints; 2) cognitive problems; 3) interpersonal sensitivity; 4) depression; 5) anxiety; 6) hostility; 7) phobic fear; 8) paranoid ideation, and 9) psychoticism. The average score of all 53 items is the BSI-Global Severity Index (BSI-GSI), which is an overall measure of psychopathology severity. In the current study the BSI-GSI and the anxiety and depression subscale (BSI-ANX and BSI-DEP) were used.

Rating scales for symptom severity of depression and anxiety
Research nurses rated the symptom severity of depression on the 10-item Montgomery Åsberg Depression Rating Scale (MADRS; (Montgomery & Åsberg, 1979)) and anxiety on the 10-item Brief Anxiety Scale (BAS; (Tyrer, Owen, & Cicchetti, 1984)). Items on both scales (e.g., "pessimistic thoughts", "worries about minor issues") are rated on a 7-point scale anchored at 4 points (1, 3, 5, and 7).

SF-36 Health Survey
Functional status was measured with the SF-36 Health Survey (Aaronson et al., 1998; Ware, Jr. & Sherbourne, 1992). The SF-36 is composed of 36 questions and standardized response choices, organized into eight multi-item scales: physical functioning (PF), role limitations due to physical health problems (RP), bodily pain (BP), general health perceptions (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and general mental health (MH). All raw scale scores are linearly converted to a 0 to 100 scale, with higher scores indicating higher levels of functioning or well-being.

Demographic characteristics
Ethnic background, education, housing situation and employment status were assessed with a self-report questionnaire. A Dutch ethnic background was assumed when the patient and both parents were born in The Netherlands.

Statistical analyses

To investigate differences between the three diagnostic groups on socio-demographic variables, chi-square tests were used on categorical variables and analyses of variance (ANOVA) on continuous variables. To test our hypotheses, we performed several ANOVAs with the severity measures as dependent variables and with diagnostic group as independent variable with 3 levels (1: pure anxiety, 2: pure depression, 3: comorbid depression and anxiety).

Subsequently, we performed separate analyses for several specific anxiety disorders: obsessive compulsive disorder (OCD), posttraumatic stress disorder (PTSD), generalized anxiety disorder (GAD), agoraphobia (AGO), panic disorder (PD) and panic disorder combined with agoraphobia (PD-AGO). For these analyses, we selected patients with singular anxiety disorders. To investigate differences between patients with- and without a comorbid depression in each of these anxiety disorders, we used t-tests with BSI-ANX as dependent variable and the presence of comorbidity as independent variable with two levels (1: only anxiety 2: anxiety and depression).

To investigate the role of age as a possible confounder, all analyses were rerun with Age as covariate. All analyses were conducted using SPSS 17. Casewise deletion was used, which resulted in different numbers for the analyses on different outcome measures.

3.3 Results

Sample characteristics

The demographic characteristics of the subjects arranged by diagnostic group are shown in Table 3.1. In the total sample (n=2278) mean age was 38.4 (SD=13.0) and 64.4% were women. A complete survey of demographic variables was available for 1919 (84%) patients in our sample. The percentage of patients with a completed survey did not differ significantly between the three diagnostic groups (resp. 85%, 85% and 83%).

No differences between the three diagnostic groups were found for gender and housing situation. The mean age did differ significantly between the three groups; the highest mean age was found in patients with a pure depression (mean age=41.0; SD=13.8). Other significant differences were found for ethnic background, educational status and employment status. Within the group 'pure anxiety' the number of patients with a Dutch ethnic background was higher than in the other two groups. Patients with comorbidity revealed a slightly lower percentage of patients with college education and a higher percentage of patients unable to work due to their sickness of disability.

	Total	Pure Anxiety	Pure Depression	Anxiety and depression	p-value
n	2278	729	860	689	
Gender (% female)	64.4	63.1	64.1	66.2	.47
Mean age (sd)	38.8 (13.0)	35.8 (12.4)$_a$	41.0 (13.8)$_b$	38.0 (12.0)$_c$	<0.001
n	1919	617	733	569	
Ethnic background (%)					.03
- Dutch	80.6	84.1	79.1	78.6	
- Other ethnicity	19.4	15.9	20.9	21.4	
Housing situation (%)					.53
- Living alone	25.7	24.5	27.6	24.6	
- Living with partner	49.9	51.1	49.4	49.2	
- Living with family	24.4	24.5	23.1	26.2	
Educational status (%)					<.001
- Lower education	10.4	7.0	11.1	13.4	
- High school (lower)	33.8	32.1	34.2	35.0	
- High school (higher)	38.4	41.8	35.5	38.3	
- College/university	17.5	19.1	19.2	13.4	
Employment status (%)					<.001
- Employed - part time	21.6	26.3	21.8	16.2	
- Employed - full time	22.0	29.3	20.3	16.2	
- Unemployed/retired	28.9	25.6	30.6	30.2	
- Unable to work due to sickness or disability	27.6	18.8	27.3	37.4	

Note. Means having a different subscript are significantly different at $p < .05$ in the Tukey difference comparison.

Table 3.1 Characteristics by diagnostic groups: gender and age and demographic variables.

Subscale mean (sd)	Pure Anxiety (n=705)	Pure Depression (n=833)	Anxiety and depression (n=663)	F (df: 2, 2198)	p-value
BSI-GSI	1.06 (0.61)$_a$	1.35 (0.66)$_b$	1.74 (0.73)$_c$	180.00	<.001
Symptoms of depression					
- BSI-DEP	1.12 (0.85)$_a$	2.00 (0.92)$_b$	2.23 (0.96)$_c$	291.32	<.001
- MADRS	13.77 (8.02)$_a$	23.30 (7.91)$_b$	25.49 (8.16)$_c$	429.39	<.001
Symptoms of anxiety					
- BSI-ANX	1.45 (0.91)$_a$	1.34 (0.87)$_a$	1.96 (0.97)$_b$	92.86	<.001
- BAS	14.32 (6.47)$_a$	14.71 (5.70)$_a$	19.05 (6.66)$_b$	123.81	<.001

Note. MADRS = Montgomery - Åsberg Depression Rating Scale; BAS = Brief Anxiety Scale; BSI = Brief Symptom Inventory. BSI-GSI = global severity index. BSI-DEP = depression subscale. BSI-ANX = anxiety subscale.
Means having a different subscript are significantly different at $p < .05$ in the Tukey difference comparison.

Table 3.2 Global symptom severity and symptoms of depression and anxiety by diagnostic group in 2201 outpatients.

Diagnostic groups

Global Severity Index

The ANOVA with the BSI-GSI as dependent variable showed a significant overall-effect of diagnostic group ($p<0.001$). The pure anxiety group had a lower mean BSI-GSI than the pure depression group; the comorbid group had a higher mean BSI-GSI than both pure disorder groups (See Table 3.2).

Depression severity

The ANOVAs with measures of depression severity (BSI-DEP and MADRS) as dependent variable showed a significant overall-effect of diagnosis ($p<0.001$). The pure anxiety group had lower mean BSI-DEP and MADRS scores than the pure depression; the comorbid group had higher mean scores than both pure disorder groups (See Table 3.2).

Anxiety severity

The ANOVAs with the two measures of anxiety severity (BSI-ANX and BAS) as dependent variable showed a significant overall-effect of diagnostic group ($p<0.001$). The comorbid group had higher mean BSI-ANX and BAS scores than

both pure disorder groups; the scores did not differ between the pure groups (See Table 3.2).

Functional severity
The ANOVAs with each of the subscales of the SF-36 as dependent variable showed an overall effect of diagnostic group on all scales ($p<0.001$). The pure anxiety group had a higher mean score than the pure depression group on all subscales (i.e. a less severe functional impairment). The comorbid group had lower mean scores (i.e. higher functional impairment) than each of the pure groups (See Table 3.3). Only on the subscale 'Role limitations due to emotional problems (RE)', the comorbid group did not have a significant lower score than the pure depression group.

Because four of the eight subscales of the SF-36 had a skewed distribution, these variables were log-transformed (PF, BP) or dichotomized (RP, RE). Subsequently, to evaluate the impact of the skewness on the results, we repeated the analyses with the transformed variables: ANOVAs with the continuous outcome variables (PF, BP) and chi-square tests in with the dichotomous outcome variables (RP, RE). The results for the transformed SF-36 scales were highly similar to the results for the untransformed scales (See Table 3.3): all group differences were significant for all subscales except for RE, and in the same direction.

Specific anxiety disorders
The mean scores on BSI-ANX were compared between specific anxiety disorders with and without comorbid depression; mean BSI-ANX scores for the different groups are shown in Figure 3.1. T-tests revealed that in all but one of the specific anxiety disorder groups, the score on the BSI-ANX was significantly higher in patients with a comorbid depression compared to patients with only an anxiety disorder. Only for patients with GAD, mean scores on the BSI-ANX did not differ between the pure GAD group and the group with GAD and depression.

Adjustment for age
When the analyses were adjusted for age, only the difference in BSI-ANX between patients with a panic disorder and agoraphobia with and without a comorbid depression was no longer significant. All other effects remained unchanged.

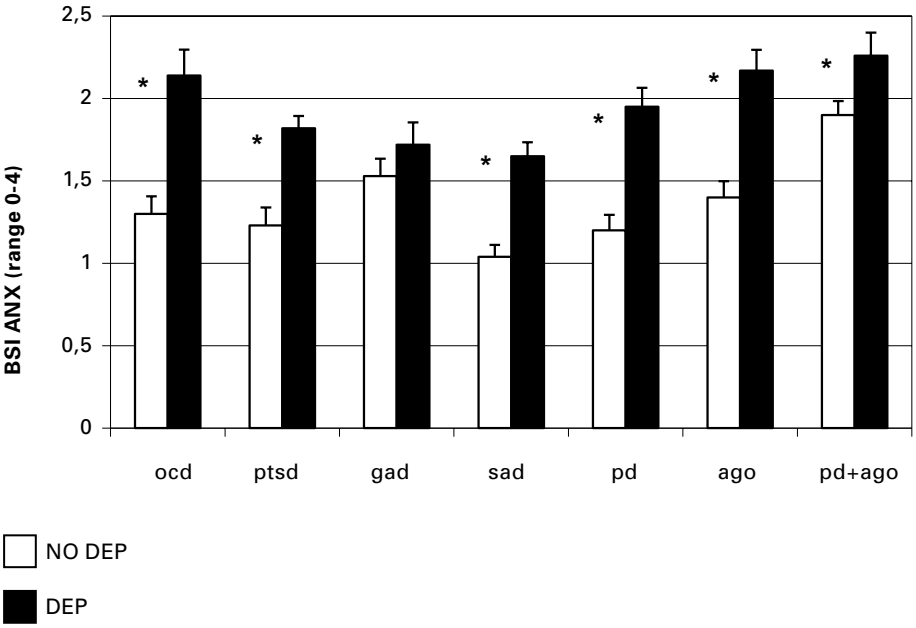
3.4 Discussion
We examined the severity of depressive and anxiety symptoms as well as functional impairment in patients with comorbid depression and anxiety disorders compared to those in pure disorders in a large general psychiatric

SF-36 subscale mean (sd)	Pure Anxiety (n=696)	Pure Depression (n=815)	Anxiety and depression (n=641)	F (df: 2, 2150)	p-value
Physical functioning (PF)	82.9 (20.5)a	73.6 (24.0)b	70.5 (24.4)c	54.06	<.001
Social functioning (SF)	55.8 (25.7)a	41.4 (25.3)b	34.8 (22.7)c	128.7	<.001
Role limitations due to physical health problems (RP)	54.5 (41.9)a	36.4 (38.7)b	30.4 (37.5)c	69.8	<.001
Role limitations due to emotional problems (RE)	41.5 (40.3)a	22.6 (32.0)b	19.3 (29.7)b	84.66	<.001
General mental health (MH)	50.3 (16.5)a	35.8 (16.0)b	31.9 (14.2)c	264.3	<.001
Vitality (VT)	45.2 (17.4)a	29.9 (15.6)b	27.5 (14.9)c	250.3	<.001
Bodily pain (BP)	75.8 (24.2)a	66.2 (27.1)b	61.5 (28.6)c	50.8	<.001
General health perception (GH)	58.9 (20.8)a	51.6 (21.1)b	47.5 (19.8)c	53.23	<.001

Note. SF-36 denotes Short Form 36 (RAND 36). Scores are on a 100 point-scale. A higher score corresponds to better functioning / health status. Means having a different subscript are significantly different at $p < .05$ in the Tukey difference comparison.

Table 3.3 Functional severity (SF-36 subscales) by diagnostic group in 2153 outpatients.

outpatient sample. The results confirmed our hypotheses and showed that in routine clinical practice, patients with comorbidity have a higher global and specific symptom severity and suffer more from severe functional impairment than patients with a pure depressive or anxiety disorder. Our main finding is that depression severity in the comorbid group was higher than that in the pure depression group, and anxiety severity in the comorbid group was higher than that of the pure anxiety group. Apparently, having an anxiety disorder in addition to a depression does not only increase the severity of anxiety symptoms, but also the severity of depressive symptoms. Similarly, having a depression in addition to an anxiety disorder does not only increase the



BSI ANX = anxiety subscale. NO DEP = no concurrent depression, DEP = concurrent depression. OCD = obsessive compulsive disorder, PTSD = posttraumatic stress disorder, GAD = generalized anxiety disorder, SAD = social anxiety disorder, PD = panic disorder, AGO = agoraphobia, PD+AGO = Panic disorder combined with agoraphobia.
*= p-value < 0.05 (t-test).

Figure 3.1 Mean (std error) scores on BSI anxiety scale for specific anxiety disorders without and with comorbidity with depression.

severity of depressive symptoms, but also the severity of anxiety symptoms. The latter was also the case when analyses were repeated for specific DSM-IV anxiety disorders.

Our findings are in line with the extended literature on comorbidity. Several studies found within a group of patients with MDD, that those with many anxiety symptoms were more severely depressed (e.g. Fava et al., 2004; Joffe, Bagby, & Levitt, 1993). Other studies focused on a specific anxiety disorder and reported higher anxiety severity or functional impairment (e.g. Wittchen et al., 2000; Cassin, Richter, Zhang, & Rector, 2009) in patients with comorbidity. We replicated these findings in a sample of patients with pure depression, pure

anxiety disorders and patients with comorbidity of depression and anxiety. With this study design, the symptom severity of the pure disorders could also be taken into account. Our data show that many patients with a 'pure' DSM IV depression also have anxiety symptoms, and patients with a 'pure' DSM IV anxiety disorder also have depressive symptoms. Moreover, we found that the mean scores on the anxiety measures (BSI-anx & BAS) did not differ significantly between patients with a pure depression and patients with a pure anxiety disorder.

Our findings are limited by the fact that we had no control group of subjects derived from the general population. This would have enabled us to determine whether the effect of diagnostic group on severity is cumulative or interactive, i.e. whether the increased symptom severity in comorbid patients equals the sum of the severity scores of the separate diagnoses or whether severity is exponentially increased when more than one diagnosis is present.

The finding that patients with comorbidity have increased anxiety severity compared to patients with a pure anxiety disorder was replicated for subgroups of patients with OCD, PTSS, PD, SAD, and Agoraphobia in our study. However, there was no difference in severity of anxiety symptoms for patients with comorbid depression in the subgroup of patients with GAD. Also, GAD revealed the highest score on depressive symptoms compared to the other anxiety disorders. This finding is in line with previous studies on the structure of DSM-IV anxiety and depression diagnoses that have shown that GAD is best grouped together with depression in a cluster of distress disorders, whereas all other anxiety disorders are grouped together in another class of fear disorders (Krueger, 1999; Watson, 2005). Our results provide further evidence that GAD might be more closely linked to depression than to other anxiety disorders.

An important strength of the current study is the large size and representativeness of the used sample. Also, the included patients were all well characterized and the sample comprised both a broad range of pure disorders and different forms of comorbidity, which enabled us to investigate the influence of comorbid depression in patients with different specific anxiety disorders. Moreover, we examined functional impairment in addition to symptom severity.

We believe that our findings give further support to the claims that depression and anxiety disorders should not be investigated in isolation (e.g. Beuke, Fischer, & McDowall, 2003). Moreover, the findings are in line with the idea that 'the use of categorical diagnostic approaches and dimensional rating scale in tandem will facilitate identification of meaningful phenotypes for future genetic, biochemical, neuroimaging, and treatment studies' (pag.73, Kaufman & Charney, 2000). This is not only relevant for research, but also for clinical practice. When using both a diagnostic interview and several severity measures at intake, a large amount of additional relevant information becomes

available to the clinician. This information can help to decide which treatment is most suitable for the patient at hand, and at follow-up, what effect the treatment has on different sorts of symptoms. Ultimately, large sets of these naturalistic data could be used to find an optimal treatment approach for patients with comorbid depression and anxiety.

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, forth 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 $a \geq 0.88$. 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)[a]				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.
[a] 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
Rating scales:				
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
Self-report:				
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 *d*
	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).

CHAPTER 5

DISTINGUISHING BETWEEN DEPRESSION AND ANXIETY: A PROPOSAL FOR AN EXTENSION OF THE TRIPARTITE MODEL

European Psychiatry 2010; 25, 197-205

Margien E. den Hollander-Gijsman
Edwin de Beurs
Nic J.A. van der Wee
Yanda R. van Rood
Frans G. Zitman

Abstract

Aim

The aim of the current study was to develop scales that assess symptoms of depression and anxiety and can adequately differentiate between depression and anxiety disorders, and also can distinguish within anxiety disorders. As point of departure, we used the tripartite model of Clark and Watson which discerns three dimensions: negative affect, positive affect and physiological hyperarousal.

Methods

Analyses were performed on the data of 1449 patients, who completed the Mood and Anxiety Symptom Questionnaire (MASQ) and the Brief Symptom Inventory (BSI). From this, 1434 patients were assessed with a standardized diagnostic interview.

Results

A model with five dimensions was found: depressed mood, lack of positive affect, somatic arousal, phobic fear and hostility. The scales appear capable to differentiate between patients with a mood and with an anxiety disorder. Within the anxiety disorders, somatic arousal was specific for patients with panic disorder. Phobic fear was associated with panic disorder, simple phobia and social anxiety disorder, but not with generalized anxiety disorder.

Conclusions

We present a five factor model as an extension of the tripartite model. Through the addition of phobic fear, anxiety is better represented than in the tripartite model. The new scales are capable to accurately differentiate between depression and anxiety disorders, as well as between several anxiety disorders.

5.1 Introduction

Anxiety and depression are highly associated. High comorbidity is repeatedly found between mood disorders and anxiety disorders (e.g. De Graaf et al., 2002; Kessler, Chiu, Demler, & Walters, 2005; Mineka et al., 1998). Some argue that both disorders are best understood as subtypes of a general neurotic syndrome (e.g. Andrews, 1996). In the most widely used diagnostic system, the DSM-IV (American Psychiatric Association, 1994), however, anxiety and depression are conceptualized as diagnostically distinct disorders. Either way, methods are needed that can adequately differentiate between depression and anxiety, in order to effectively study their relation and their (shared or distinct) etiological factors.

An important contribution to the field was made by Clark and Watson who introduced the tripartite model to assess distinctive and overlapping features of mood disorders and anxiety (1991). To make a model that is able to differentiate between patients with a depression and patients with an anxiety disorder, they took the basic dimensions of affect, *negative affect* and *positive affect,* and added a third factor *physiological hyperarousal*. Thus, their model comprises three dimensions to describe the symptomatology of depression and anxiety. *Negative affect* consists of symptoms of general distress and is common to both depression and anxiety disorders. *Positive affect* is referring to enthusiasm, excitement and energy. A low score on this dimension is typical for depression. The third dimension *physiological hyperarousal,* consist of somatic tension and arousal and was presented as specific for anxiety. To operationalize and measure the dimensions of the tripartite model, Watson and Clark developed the Mood & Anxiety Symptom Questionnaire (MASQ, Watson & Clark, 1991).

The tripartite model has inspired a large body of research, but has met some criticism as well. We will review the major points of critique on each of the three factors of the tripartite model and the MASQ. First, there is criticism on the general distress factor: negative affect (NA) comprises a heterogeneous group of symptoms such as anger, fear, and tension. Although it is supposed to be non-specific for depression as well as for anxiety, some symptoms are in fact quite specific for either depression or anxiety disorders. Anger attacks, for instance, are twice as prevalent in depressed patients as in patients with anxiety disorders ((Pasquini, Picardi, Biondi, Gaetano, & Morosini, 2004; Picardi, Morosini, Gaetano, Pasquini, & Biondi, 2004), Gould et al. 1996 in (Shankman & Klein, 2003)). Symptoms like 'felt afraid' and 'felt nervous' on the other hand, seem more specific for anxiety disorders than for depression. Clark and Watson acknowledged the heterogeneity of NA by subdividing the general distress items of the MASQ in three subcategories: General Distress Depression (GDD), General Distress Anxiety (GDA) and General Distress Mixed (GDM). The items were assigned to these subcategories on face value, i.e. on their similarity with DSM-IIIR criteria for either depression or anxiety

disorders. The separateness of these three subscales of NA, however, could not be confirmed with factor analyses: Several studies on the MASQ concluded that a two or three factor model had the best fit (Boschen & Oei, 2006; Buckby et al., 2008; Burns & Eidelson, 1998; De Beurs et al., 2007; Keogh & Reidy, 2000; Watson et al., 1995).

Second, there are criticisms on the scale positive affect (PA), called anhedonic depression (AD) in the MASQ. Although the scale is presented as a single dimension with two endpoints, (a low positive affect endpoint, loss of interest and a high positive affect endpoint, feeling good), the two endpoints appear as separate factors in factor-analysis (Keogh & Reidy, 2000; Watson et al., 1995; Nitschke et al., 2001; De Beurs et al., 2007) suggesting these two endpoints represent different dimensions. Other evidence that 'lack of interest' and 'feeling good' are separate constructs can be found in the research of Tomarken and Dichter (Tomarken, Dichter, Freid, Addington, & Shelton, 2004). They reported medication to have a differential effect on 'lack of interest' and 'feeling good' as measured with the MASQ. Items assessing the low positive affect pole of the anhedonia dimension were more sensitive to earlier/lower dose bupropion SR (sustained-release) treatment, whereas items assessing the high positive affect pole were more sensitive to later/higher dose bupropion SR treatment.

The third factor, physiological hyperarousal, has been criticised as well. Firstly, a high score on the MASQ scale *'anxious arousal'* (AA) seems to be characteristic for panic disorder only and not for all anxiety disorders (e.g.(Mineka et al., 1998; Chorpita, 2002; Joiner et al., 1999)). Somatic signs of anxiety dominate this scale, while other anxiety symptoms are not well represented. In addition, the scale does not distinguish between patients with and without an anxiety disorder (Buckby, Yung, Cosgrave, & Cotton, 2007). It also does not distinguish between patients with anxiety disorder and depression (Boschen & Oei, 2007; De Beurs et al., 2007). Buckby (Buckby et al., 2007) even found significant higher scores for depressed patients over anxious patients on AA.

The possible limitations of the tripartite model have been acknowledged by the original authors. They recommend to view in future research „individual disorders as representing unique *combinations* of different types of symptoms, with each type showing varying degrees of nonspecificity and with no type being entirely unique to any single disorder" (Mineka et al., 1998, p.398). To operationalize this idea, adequate symptom scales must be developed to include the more unique symptoms of specific mood- and anxiety disorders in addition to common symptom scales. In such a dimensional approach to psychopathology, every disorder (and every patient) will have a more or less unique profile. This is a valuable recommendation and underlines the need for scales that can represent more adequately relevant aspects of anxiety.

The aim of the current study was to develop scales that assess symptoms of depression and anxiety and can adequately differentiate between depression and anxiety disorders, and also can distinguish within anxiety disorders. First, we explored the factor structure of the items of the MASQ extended with items of a questionnaire containing many anxiety items. We choose the Brief Symptom Inventory (BSI), because this questionnaire has many anxiety-related items and resembles the MASQ in the construction of the items and time span. We expected to find in addition to the two specific scales of the MASQ (AA and AD) at least one extra scale, based to a large extent on BSI-items, tapping other aspects of anxiety than the AA scale. We expected the set of new scales to cover anxiety more adequately than the tripartite model and to provide profiles specific for depression and each of the anxiety disorders.

To examine the psychometric properties of the scales found with factor-analysis, we addressed the following questions: (a) What is the reliability of the scales based on the found factor structure? (b) Do these scales measure more distinct concepts as compared to the original scales of the MASQ? (c) Do the new scales have good discriminatory validity for depression and anxiety? (d) Are the new scales able to differentiate between specific anxiety disorders? We expected the set of new scales to have a good reliability and better discriminatory validity than the original MASQ scales.

5.2 Materials and methods

Study sample

This study was conducted on data collected through Routine Outcome Monitoring (ROM, (De Beurs et al., 2011)). ROM is a monitoring system for patient care, implemented in the outpatient clinics of Rivierduinen (a large organization for the provision of mental health care in the province of Zuid-Holland, the Netherlands) and the department of psychiatry of the Leiden University Medical Center (LUMC). All patients referred to these clinics for treatment of a mood-, anxiety- or somatoform disorder, have an assessment session with a psychiatric research nurse at the start, during, and at the end of the treatment. During the first session, a standardized diagnostic interview is administered and interviewer and self-reported ratings are determined. The sample consisted of 1479 patients admitted consecutively between January 2002 and March 2005 to the outpatient clinics of the Rivierduinen Psychiatric Hospital (754 in Leiden; 198 in Alphen a/d Rijn, 163 in Leidschendam and 163 in Voorhout) and the psychiatric outpatient department of Leiden University Medical Center (LUMC, n=201).

Measures

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

The M.I.N.I. is a short clinical diagnostic interview developed to explore the

presence of current and life-time Axis-I disorders according to the DSM-IV diagnostic criteria (Sheehan et al., 1998). The Dutch translation of the M.I.N.I. Plus 5.0.0-R was used in the present study (Van Vliet et al., 2000). The M.I.N.I.-Plus is an extended version of the original M.I.N.I.. Lecrubier and colleagues (Lecrubier et al., 1997) report sufficient reliability of the M.I.N.I; Inter-rater reliability ranged from k=.88 to 1.00, test-retest reliability ranged from 0.76 to 0.93, validity was demonstrated by sufficient concordance with the Composite International Diagnostic Interview (CIDI, WHO). Psychiatric research nurses who were extensively trained and supervised performed the interviews. All diagnoses reported in this paper were current at the time of investigation.

Mood and Anxiety Symptom Questionnaire (MASQ, (Watson & Clark, 1991)).
The MASQ consists of 90 items, allocated to five subscales: 1) anhedonic depression; 2) anxious arousal; 3) general distress depression; 4) general distress anxiety, and 5) general distress mixed. All items are presented with a five-point rating scale ranging from 1 (not at all) to 5 (very much). We used a Dutch adaptation of the MASQ (De Beurs et al., 2007).

Brief Symptom Inventory (BSI)
The Brief Symptom Inventory (Derogatis & Melisaratos, 1983; De Beurs, 2005) is a shortened version of the Symptom Checklist (SCL-90) (Arrindel & Ettema, 1986; Derogatis et al., 1973), and is used to measure psychological complaints or symptoms. The BSI consists of 53 items that are rated on a five-point Likert type scale, ranging from 0 (not at all) to 4 (very much). The items are assigned to nine dimensions: 1) somatic complaints; 2) cognitive problems; 3) interpersonal sensitivity; 4) depression; 5) anxiety; 6) hostility; 7) phobic fear; 8) paranoid ideation, and 9) psychoticism.

Pool of items.
For this study a selection of items was made from the BSI and the MASQ. From the MASQ the 77 of 90 items which were assigned to a subscale by the authors of the MASQ (Watson & Clark, 1991) were used. The BSI subscales *paranoid ideation* and *psychoticism* were not used, because we are predominantly interested in mood and anxiety disorders and expected a high positive skewness on these items in the population we studied. Twelve items of the BSI closely resembling MASQ items were omitted, as highly collinear items should not be subjected to factor-analysis. The end result was a pool of 104 items.

Statistical analyses
All positively formulated items of the MASQ were reversed keyed before analysis. Exploratory factor analyses with oblique rotation were performed using SPSS procedure 'Factor, rotations Oblimin'. We preferred factor analysis

(FA) over principal component analysis (PCA), because we were looking for factors, uncorrelated or correlated, which explain the interrelationships between the observed variables. This approach is different from the idea of PCA in which uncorrelated components are sought which explain the most variation of the variables. Oblique rotation rather than orthogonal rotation was chosen, because substantial correlation between the factors was expected. Before factor extraction, the correlations between the items were inspected to check for items which failed to correlate 0.20 or more with any other item (Floyd & Widaman, 1995). None were found.

The number of factors to extract was determined using eigenvalues above one, a parallel analysis (Monte Carlo PCA), the screeplot, the number of unique loading items per factor, and most importantly, the interpretability of the factors. New subscales (mean score) were formed with items loading at least .40 on the factors (loadings in the pattern matrix). Before calculating the scores on the new subscales all items of the BSI were recoded from 0-4 to 1-5 to match the scores of the MASQ.

After new subscales were composed, the reliability and validity of these scales were determined. Coefficient α was used to assess the internal consistency of the scales (question a). To examine the level of distinctiveness of the new scales (question b), correlation coefficients (Pearson's R) between all scales were calculated. To determine the discriminant validity of the newly found scales, we investigated whether subscale scores could discriminate between subgroups of patients based on diagnostic information obtained with the M.I.N.I.-Plus. A stepwise discriminant function analysis was performed to investigate the ability of the new scales to discriminate between the two diagnostically purest groups: 1) patients with one or more anxiety disorder(s) but without a depression and 2) patients with a depression but without an anxiety disorder (question c).

To determine whether different anxiety disorders reveal a different symptom profile (question d), we compared the mean scores of groups of patients with different anxiety disorders (and no comorbid depression) with a multivariate analysis of variance (MANOVA). All analyses were conducted using SPSS-16.

5.3 Results

Sample description

The mean age was 39 years, (sd=13, range 18 – 82) and 936 patients (63.3%) were female. All patients (n=1479) completed the MASQ and 1449 (98%) also the BSI. The M.I.N.I.-PLUS was administered in 1434 (97%) patients. Criteria for at least one current Axis-I DSM-IV disorder were met by 1347 patients (94%) and for at least two current disorders by 947 (64%) patients: mood-, anxiety- and somatoform disorders were diagnosed in 52%, 57% and 21% of the patients respectively. Depression includes both depression (89%) and dysthymia (11%),

item	questionnaire	nr
Felt like a failure	masq	47
Felt worthless	masq	13
Felt inferior to others	masq	64
Was disappointed in myself	masq	74
Blamed myself for a lot of things	masq	24
Felt hopeless	masq	22
Felt withdrawn from other people	masq	26
Felt unattractive.	masq	53
Feeling hopeless about the future	BSI	35
Felt discouraged	masq	8
Felt pessimistic about the future	masq	42
Feeling very self-conscious with others	BSI	42
Felt dissatisfied with everything	masq	29
Felt depressed	masq	16
Felt uneasy	masq	20
Feeling lonely	BSI	16
Felt like nothing was very enjoyable	masq	33
Had trouble making decisions	masq	70
Feeling blue	BSI	17
Felt sad	masq	6
Felt like there wasn't anything interesting or fun to do	masq	44
Feeling that people are unfriendly or dislike you	BSI	21
Thought about death or suicide	masq	89
Feeling blocked in getting things done	BSI	15
Worried a lot about things	masq	84
Felt like I was having a lot of fun	masq	23
Felt optimistic	masq	18
Felt like I had a lot to look forward to	masq	40
Looked forward to things with enjoyment	masq	30
Felt really "up" or lively	masq	58
Felt like I had accomplished a lot	masq	35
Felt like I had a lot of interesting things to do	masq	36
Felt really good about myself	masq	86
Felt really happy	masq	14
Felt cheerful	masq	1
Was proud of myself	masq	49
Felt hopeful about the future	masq	78

Table 5.1 Factor structure after OBLIMIN rotation.

original scale	DM	LPA	SA	PF	HOS
GDD	**0.84**				
GDD	**0.82**				
GDD	**0.79**				
GDD	**0.77**				
GDD	**0.67**				
GDD	**0.66**				
AD (interest)	**0.63**				
AD (interest)	**0.61**				
DEP	**0.60**				
GDD	**0.60**				
GDD	**0.60**				
I-S	**0.55**				
GDM	**0.54**				
GDD	**0.51**				
GDA	**0.51**				
DEP	**0.51**				
AD (interest)	**0.49**	-0.31			
GDM	**0.49**				
DEP	**0.47**	-0.31			
GDD	**0.44**				
AD (interest)	**0.44**				
I-S	**0.43**				0.32
AD (interest)	**0.42**				
O-C	**0.41**				
GDM	**0.40**				
AD (PA)		**-0.86**			
AD (PA)		**-0.84**			
AD (PA)		**-0.82**			
AD (PA)		**-0.78**			
AD (PA)		**-0.77**			
AD (PA)		**-0.77**			
AD (PA)		**-0.75**			
AD (PA)		**-0.73**			
AD (PA)		**-0.72**			
AD (PA)		**-0.70**			
AD (PA)		**-0.68**			
AD (PA)		**-0.67**			

item	questionnaire	nr
Felt like I had a lot of energy	masq	72
Seemed to move quickly and easily	masq	27
Felt really slowed down	masq	66
Felt faint	masq	19
Was trembling or shaking	masq	79
Muscles twitched or trembled	masq	69
Felt dizzy or light-headed	masq	52
Hands were shaky	masq	57
Heart was racing or pounding	masq	75
Muscles were tense or sore	masq	81
Trouble getting your breath	BSI	29
Felt numbness or tingling in my body	masq	25
Was short of breath	masq	55
Had pain in my chest	masq	45
Had hot or cold spells	masq	48
Had trouble swallowing	masq	87
Hands were cold or sweaty	masq	88
Felt nauseous	masq	9
Feeling weak in parts of your body	BSI	37
Felt like I was choking	masq	61
Had a lump in my throat	masq	65
Had a very dry mouth	masq	67
Had an upset stomach	masq	63
Got tired or fatigued easily	masq	90
Having to avoid certain things, places, or activities because they frighten you	BSI	31
Suddenly scared for no reason	BSI	12
Spells of terror of panic	BSI	45
Feeling afraid to travel on buses, subways, or trains	BSI	28
Feeling afraid in open spaces or on the streets	BSI	8
Your feelings being easily hurt	BSI	20
Temper outburst that you could not control	BSI	13
Feeling easily annoyed or irritated	BSI	6
Getting into frequent arguments	BSI	46
Having urges to break or smash things	BSI	41
Felt irritable	masq	17
Having urges to beat, injure, or harm someone	BSI	40
Feeling tense or keyed up	BSI	38

Table 5.1 Factor structure after OBLIMIN rotation (Continued).

original scale	DM	LPA	SA	PF	HOS
AD (PA)		**-0.64**			
AD (PA)		**-0.51**			
AD (interest)	0.35		**0.48**		
AA	0.30		**0.45**	-0.32	
AA			**0.72**		
AA			**0.71**		
AA			**0.67**		
AA			**0.63**		
AA			**0.61**		
GDA			**0.60**		
SOM			**0.60**		
AA			**0.60**		
AA			**0.58**		
AA			**0.56**		
AA			**0.56**		
AA			**0.55**		
AA			**0.53**		
GDA			**0.51**		
SOM			**0.50**		
AA			**0.50**		
GDA			**0.48**		
AA			**0.47**		
GDA			**0.46**		
GDM			**0.44**	-0.37	
PHOB	0.31			**0.45**	
ANX			0.37	**0.48**	
ANX			0.35	**0.50**	
PHOB			0.30	**0.44**	
PHOB				**0.44**	
I-S	0.32				**0.49**
HOS					**0.79**
HOS					**0.76**
HOS					**0.75**
HOS					**0.72**
GDM					**0.66**
HOS					**0.60**
ANX					**0.54**

item	questionnaire	nr
Felt keyed up, “on edge”	masq	82
Felt confused	masq	4
Felt sluggish or tired	masq	56
Feeling uneasy in crowds, such as shopping or at a movie	BSI	43
Feeling no interest in things	BSI	18
Felt really bored	masq	21
Had trouble paying attention	masq	80
Felt like something awful was going to happen	masq	34
Felt like it took extra effort to get started	masq	39
Felt afraid	masq	2
Slept very well	masq	5
Was afraid I was going to die	masq	73
Startled easily	masq	3
Nervousness or shakiness inside	BSI	1
Felt nervous	masq	15
Felt very restless	masq	50
Had to urinate frequently	masq	85
Feeling nervous when you are left alone	BSI	47
Had trouble concentrating	masq	76
Feeling so restless you couldn’t sit still	BSI	49
Having to check and double-check what you do	BSI	26
Had diarrhea	masq	12
Had trouble remembering things	masq	31
Your mind going blank	BSI	32
Did not have much of an appetite	masq	37
Was unable to relax	masq	59
Had trouble falling asleep	masq	51
Had trouble staying asleep	masq	83
Felt like crying	masq	10
Felt tense or “high-strung”	masq	77

Note:
Original scales: MASQ: AA = anxious arousal, AD = anhedonic depression, GDD = general distress depression, GDA = general distress anxiety, GDM = general distress mixed. BSI: SOM = Somatization, O-C = Obsessive-Compulsive, I-S = Interpersonal Sensitivity, DEP = Depression, ANX = Anxiety, HOS = Hostility, PHOB = Phobic Anxiety.

Table 5.1 Factor structure after OBLIMIN rotation (Continued).

original scale	DM	LPA	SA	PF	HOS
GDA					**0.51**
GDM	0.39				
GDD	0.38		0.36	-0.36	
PHOB	0.38			0.35	
DEP	0.37	-0.30			
AD (interest)	0.37				
GDM	0.33			-0.31	
GDM	0.31		0.32		
AD (interest)	0.30		0.36	-0.35	
GDA	0.30			0.39	
GDM		-0.35			
AA			0.39		
AA			0.38		
ANX			0.37		
GDA			0.31		
GDM			0.31		
AA			0.30		
PHOB				0.32	
GDM				-0.31	
ANX					0.36
O-C					0.35
GDA					
GDM					
O-C					
GDM					
GDA					
GDM					
GDM					
GDD					
GDA					

New scales: DM = depressed mood, LPA = low positive affect, SA = somatic arousal, PF = phobic fear, HOS = hostility.

but not bipolar disorder. The comorbidity between anxiety and depression was high: 30.1% of the patients had a depression as well as an anxiety disorder. The percentage of patients with one or more anxiety disorder(s) and no depression was 26.8 and the percentage of patients with a depression and no anxiety disorder was 21.5. The remaining group of patients (21.5%) had no anxiety or depression diagnoses and consisted of patients with a somatoform disorder or an adjustment disorder or no current disorder (6%) according to M.I.N.I.-Plus.

Factor Analysis

The 104 items were subjected to an exploratory factor analysis using the data of patients who completed the MASQ and BSI (n=1449). Parallel analysis suggested retaining 10 factors. Sixteen factors had an eigenvalue > 1 and the eigenvalues of the first 6 factors were: 33.8, 6.6, 4.4, 3.0, 2.5 and 2.0. The screeplot of the factor solution flattened out starting from the fifth or sixth component, suggesting that a four or five-factor solution would best fit the data (Catell, 1966). Rotated factor solutions (oblimin rotation) were calculated for three, four-, and five-factor solutions. The three factor solution resembled the tripartite model and explained 43% of the variance. Most items with uniquely loadings, loaded on the first factor (e.g. sad, angry, low self-esteem, guilty, unattractive, and worrying). The second factor contained all the positively skewed items (PA), and only items with explicit physical symptoms loaded uniquely on the third factor. The four factor solution (explaining 46% of the variance) resembled the three factor solution supplemented with a factor with only one item with a unique loading ("Having to avoid certain things, places, or activities because they frighten you" (BSI item 31)). In the five-factor solution no items had a loading higher than .399 on more than one factor (see Table 5.1) making this solution easier to interpret. Thus, a five-factor solution was chosen. The five-factor solution accounted for 48% of the total variance in the scores. The first factor loads mainly on items originally belonging to two subscales of the MASQ: general distress depression and the negative endpoint of anhedonic depression: loss of interest. The second factor represents the positive endpoint of the MASQ subscale anhedonic depression: positive affect. The items loading high on the third factor are mainly items about somatic symptoms of anxiety. The fourth factor is a combination of items of the BSI subscales anxiety and phobic anxiety. The fifth factor comprises predominantly items from the BSI subscale hostility. When comparing our five factors with the five scales of the MASQ, the scales anhedonic depression (low positive affect) and anxious arousal are retained. Three new scales emerged: depressed mood, phobic fear and hostility.

Subsequently, five scales were composed by calculating the mean of the items with loadings of at least .40 (the items in bold typeface in Table 5.1): depressed mood (DM; factor I), low positive affect (LPA, factor II), somatic arousal (SA;

	DM	LPA	SA	PF	HOS
DM (26 items)	(α= .96)	.64	.57	.48	.68
LPA (14 items)		(α= .94)	.41	.26	.39
SA (23 items)			(α= .93)	.54	.54
PF (5 items)				(α= .85)	.45
HOS (9 items)					(α=.89)

Note:
New scales: DM=depressed mood, LPA=low positive affect, SA=somatic arousal, PF=phobic fear, HOS=hostility. All correlations are significant at the 0.01 level.

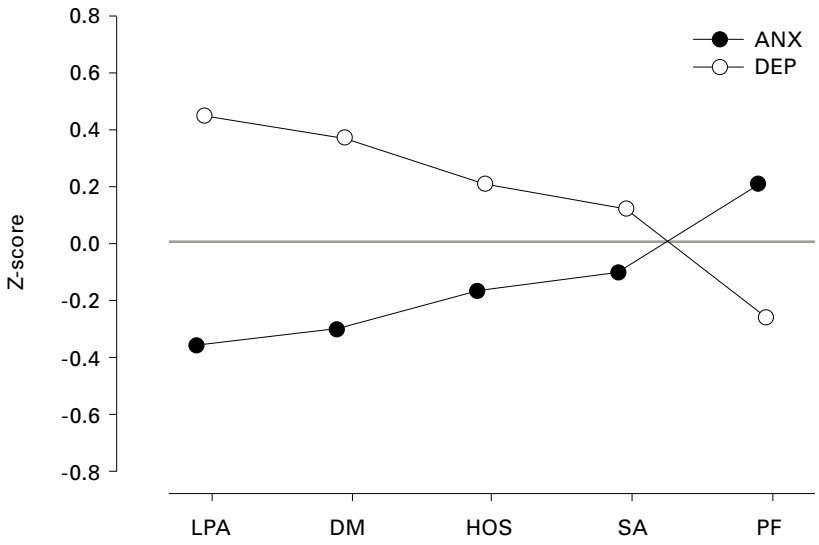
Table 5.2 Correlation & Cronbach's alpha among newfound scales.

	AD	AA	GDD	GDA	GDM
AD	(α= .94)	.50	.79	.61	.76
AA		(α= .90)	.54	.76	.66
GDD			(α= .92)	.71	.79
GDA				(α= .85)	.77
GDM					(α= .89)

Note:
Original scales: MASQ: AD=anhedonic depression, AA=anxious arousal, GDD=general distress depression, GDA=general distress anxiety, GDM=general distress mixed. All correlations are significant at the 0.01 level.

Table 5.3 Correlation & Cronbach's alpha among original MASQ scales.

factor III), phobic fear (PF; factor IV) and hostility (HOS; factor V). Although one of the scales (PF) consists of only five items, the internal consistency (reliability) of all the scales is high, ranging between α=.85 and α=.96 (depicted on the diagonal in Table 5.2). Furthermore, the new scales appear sufficiently distinct: The correlations among the new scales (Table 5.2) range from .26 to .68, while those among the original scales of the MASQ (Table 5.3) range from .50 to .79. Especially low is the correlation between *low positive affect* and *phobic fear*: r=.26, suggesting that these scales may differentiate well between depression and anxiety disorders.



Note:
New scales: LPA=low positive affect, DM=depressed mood, HOS=hostility SA=somatic arousal, PF=phobic fear, ANX= patients with one or more anxiety disorder(s) and no mood disorder, DEP= patients with a mood disorder and no anxiety disorder.

Figure 5.1 Symptom profiles (mean z-scores) for patients with a singular depression (n=309) and for patients with a singular anxiety disorder (n=385).

Discriminant analysis

To determine the discriminant validity of the new scales, we made two groups of patients: 1) patients with one or more anxiety disorder(s) without a comorbid depression (ANX) and 2) patients with a depression without a comorbid anxiety disorder (DEP). Because the discriminant validity of the scales is best tested in a comparison of pure anxiety with pure depression, we did not include patients with comorbid anxiety disorder and depression in this analysis. With excluding this group, we also limited the influence of the severity of psychopathology. A stepwise discriminant function analysis was performed to investigate the ability of the five new scales to discriminate between these two diagnostically purest groups. The analysis resulted in a model ($\chi^2(3)=215$, $p \leq .001$) based on three of the five scales: low positive affect, depressed mood and phobic fear. Low positive affect and phobic fear are the best discriminators between depressed patients and patients with an anxiety disorder. Patients with a high score on

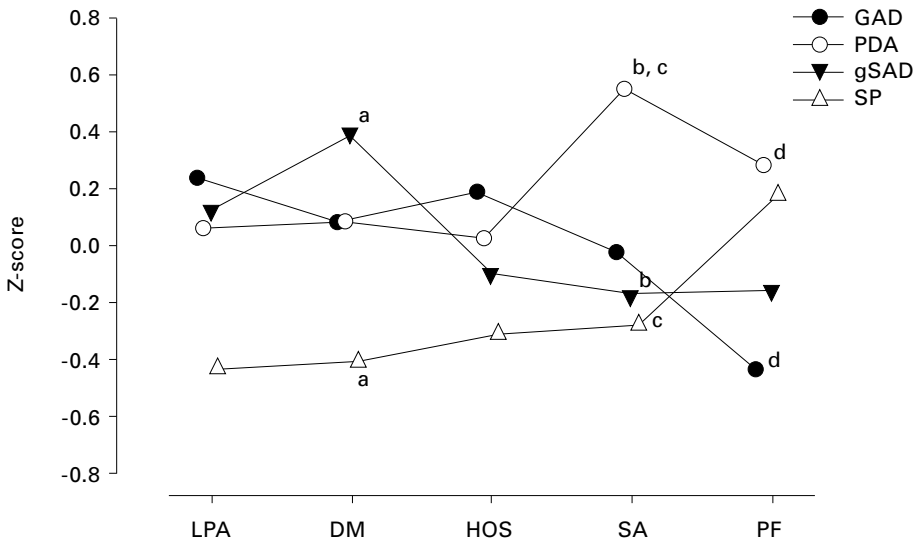
low positive affect are more likely to belong to the group with depression (mean score DEP=4.36, SD=.64 versus mean score ANX=3.70, SD=.82) and patients with a high score on *phobic fear* are more likely to belong to the group with an anxiety disorder (mean score DEP=0.60, SD=.73 versus mean score ANX=1.00, SD=.92). With this model 75% of the patients with an anxiety disorder and 71% of the patients with a depression were classified correctly. When analysed with the original MASQ these percentages are 77% and 69%, respectively. However, with the original MASQ none of the scales in the discriminant function reveal a high score specific for anxious patients: depressed patients score higher on all original MASQ-scales in the discriminant function than anxious patients.

To illustrate the discriminant value of each new scale, the results of the two groups in the discriminant analysis (patient with a depression and no anxiety disorder, and patients with an anxiety disorder and no depression) on the five scales are shown on Fig. 5.1. To make comparisons between the scales easier, we standardized scores for the five new subscales (z-scores). The figure shows that both groups have a specific profile.

If scales are specific for either depression or anxiety, one also expects that patients with the specific disorder score higher on the scale than patients without any mood or anxiety disorder (NO ANX/DEP). We tested this with two t-tests on the new scales specific for anxiety (PF) and specific for depression (LPA). As expected, patients with a depression had a higher score on LPA than patients without a mood or anxiety disorder (mean score DEP=4.36, SD=.64 versus mean score NO ANX/DEP=3.60, SD=.79, p=.000) and patients with an anxiety disorder had a higher score on PF than patients without a mood or anxiety disorder (mean score ANX=1.00, SD=.92 versus mean score NO ANX/DEP=0.37, SD=.55, p=.000).

Multivariate analysis of variance

To further examine the discriminant validity of the new found scales, we compared the scores of four groups of patients with a specific anxiety disorder. For this analysis, we only selected patients with a singular anxiety disorder (no comorbid depression and no more than one anxiety disorder). Because of the high comorbidity in our sample, this resulted in relatively small groups of patients with singular anxiety disorders: general anxiety disorder (GAD; n=32), panic disorder (PDA; n=28), simple phobia (SP; n=20) or generalized social anxiety disorder (gSAD; n=47). The mean z-scores of these four groups of patients are shown on Fig. 5.2. The figure clearly shows that patients with panic disorder have a higher score on somatic arousal than patients with any of the other anxiety disorders. Differences between the groups were tested with a multivariate analysis of variance (MANOVA). Main effects were found on *somatic arousal*, *phobic fear* and *depressed mood* (SA ($F(3)=4.4$, $p\leq.006$), PF ($F(3)=3.5$, $p\leq.019$) and DM ($F(3)=3.1$ $p\leq.029$)). Post-hoc analyses (Tukey)



Note:
New scales: LPA=low positive affect, DM=depressed mood, HOS=hostility SA=somatic arousal, PF=phobic fear, PDA=panic disorder, GAD=generalized anxiety disorder, gSAD=generalized social anxiety disorder, SP=simple phobia. Corresponding letters (a, b, c, d) refer to significant difference according to the post-hoc analyses (Tukey).

Figure 5.2 Symptom profiles (mean z-scores) for four specific anxiety disorders without a mood disorder: Panic disorder (n=28), generalized anxiety disorder (n=32), simple phobia (n=20) and generalized social anxiety disorder (n=47).

showed that the panic disorder group had higher scores than the GAD and the gSAD patients on *somatic anxiety* (SA; PDA>GAD/gSAD). Patients with a panic disorder revealed a higher score on *phobic fear* than GAD patients (PF; PDA>GAD). On the scale *depressed mood* patients with gSAD had a higher score than patients with a simple phobia (DM; gSAD>SP).

5.4 Discussion

The aim of the present study was to develop scales that can adequately differentiate between depression and anxiety disorders, and also can distinguish within the anxiety disorders. The scales are based on a pool of items from the MASQ supplemented with BSI items. Factor-analysis resulted in a solution with

five factors. The requirement of distinctness of the new scales was largely met: the correlations between the scales based on this factor solution were relatively low and the internal consistency of the scales was good. The intercorrelation of two scales, *phobic fear* and *low positive affect,* was even very modest (r=.26). In addition, the scales were able to differentiate rather well between patients with a mood and patients with an anxiety disorder. As compared to the original MASQ, the percentages correctly classified patients are highly similar. However, similar to a previous study in a clinical sample (Buckby et al., 2007), we found that depressed patients score higher on all original MASQ-scales in the discriminant function than anxious patients. The discriminant validity of the new scales was considerably better, as patients with an anxiety disorder scored significantly higher on one of the scales (PF) than patients with a depression.

Mineka, Watson and Clark (Mineka et al., 1998) suggested the use of a model in which each individual syndrome contains both a common and a unique component: the integrative hierarchical model. The scales we found fit well in this model. Patients with only an anxiety disorder (except patients with GAD) are predominantly characterized by a heightened score on *phobic fear* whereas depressed patients show high scores on the *low positive affect* and *depressed mood* scales. Furthermore, each of the four included anxiety disorders had a different profile on the five new scales. Comparing the mean scores on the new scales between patients with one of four anxiety disorders (gSAD, GAD, SP and PDA), the scale *somatic arousal* appeared to be specific for patients with panic disorder. This is consistent with the findings of several previous studies (Chorpita, 2002; De Beurs et al., 2007; Keogh & Reidy, 2000). The phenomenology of anxiety disorders is better represented with *phobic fear* next to *somatic arousal*.

The scale *hostility* did not contribute to the discriminant function. A possible explanation can lie in the recent discovery that irritability within a depression, is associated with greater overall severity, anxiety comorbidity and suicidality (Perlis et al., 2009) and therefore not specific for depression nor anxiety. However, hostility is clinical relevant and underestimated in our current classification systems (Pasquini et al., 2004; Picardi et al., 2004) and can therefore be a valuable feature to assess.

Our findings are in line with the structure underlying mood and anxiety disorders that was recently presented by Watson (Watson, 2005). For the DSM-5, Watson has suggested to use a quantitative hierarchical model in which the mood and anxiety disorders are taken together to form an overarching class of emotional disorders with 3 subclasses (distress disorders, fear disorders and bipolar disorders). Our finding that patients with GAD, just like patients with a depression do not have a high score on the two 'anxiety-like dimensions' PF and SA, is in line with the suggestion of Watson to classify GAD as a distress disorder rather than as a fear disorder.

In the present study, the content validity of the new scales was investigated by comparing scores of patients with singular disorders. Of course, such an approach is not a definite test of the validity of the proposed scales for psychopathology of depression and anxiety disorders, given the inherent limitations of diagnostic categorization by itself and the resulting overlap in the phenomenology of depression and anxiety disorders. In future research the content validity should also be evaluated in other ways as well, for instance by demonstrating a distinct predictive value of dimensions for the course of complaints over time or a prognostic value for treatment effect.

Strength of the study is the large patient sample, mainly consisting of patients with the relevant disorders: depression and anxiety disorders. We choose to limit the analyses of the discriminant validity to only those patients with a pure depressive or anxiety disorder. This was feasible given the large number of patients included in the study. An important advantage of this approach is the diminished role of severity in the analysis.

A limitation of the findings regarding the factor structure is that the subscale *low positive affect* is only composed of positively formulated items (e.g. "I felt cheerful"). The assumption of the tripartite model that the dimension positive affect comprises two extremes (lack of interest and feeling good) is thus not confirmed by our results. This finding is consistent with earlier studies that showed lack of interest as belonging to negative affect (De Beurs et al., 2007; Keogh & Reidy, 2000; Watson et al., 1995). However, the fact that all positively formulated items load predominantly on a single factor suggests a method effect, rather than the presence of a conceptually distinct construct.

In sum, we present a five factor model as an extension of the tripartite model. Through the addition of phobic fear, anxiety is better represented than in the tripartite model. The new scales are capable to accurately differentiate between depression and anxiety disorders, as well as between several anxiety disorders.

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 $\geq$ 0.90 and a RMSEA $\leq$ 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

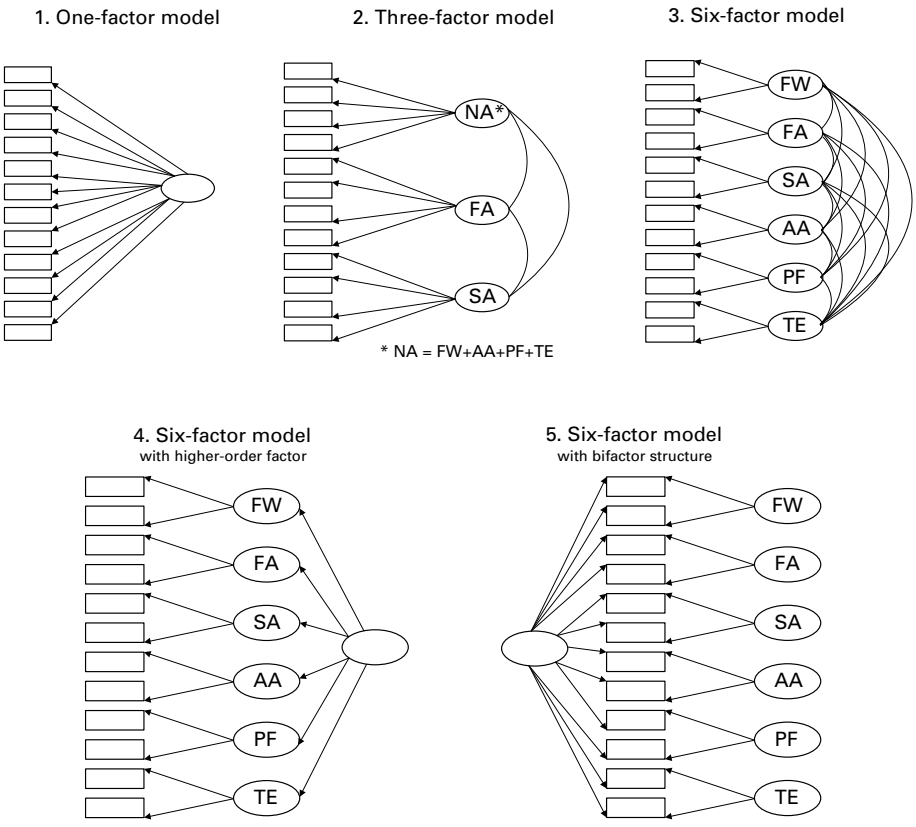


Figure 6.1 Schematic illustration of a priori structural models. The 5 pictures depict only a few relevant parameters; the exact number of items and error terms on symptoms and intermediate factors are omitted for clarity.
NA=negative affect, FW=feelings of worthlessness, FA=fatigue, SA=somatic arousal, 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 χ2	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 χ2 = 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.

CHAPTER 7

SUMMARY, DISCUSSION AND FUTURE PERSPECTIVES

7.1 Introduction

There are two approaches to psychopathology, a categorical and a dimensional approach. In the categorical approach each psychiatric disorder is characterized by a set of criteria. Diagnoses are made by checking whether a patient meets the criteria for one or more disorders. This is basically a dichotomous decision process; a patient meets criteria and therefore has a disorder or not. At first sight, this approach seems pretty straightforward for research and clinical practice. In etiological research it enables the study of well-defined patient groups. In clinical practice, the psychiatrist or therapist who made a diagnosis knows which treatments are appropriate, based on randomized controlled trials carried out in patients with the same disorder. However, patients with the same disorder may differ very much in symptomatology. For instance, if both patients have a depression (i.e. both have the required 5 out of 9 symptoms) they may have only a single symptom in common. Besides, in clinical practice comorbidity is the rule rather than the exception and this contributes to the heterogeneity. For research into the etiology of psychopathology this heterogeneity may explain why the results are often inconsistent. For clinical practice this implies that the evidence on which treatments are based is oversimplified. This may play a role in the often modest treatment results.

The dimensional view has the potential to overcome these problems as it allows a more comprehensive analysis of psychopathology. First of all, subjects are assessed not on a single, but on multiple dimensions. It is in fact a multidimensional approach. Each dimension is not assessed as present or absent, but is quantified along a continuum. And each patient is assessed along all the dimensions included in the investigation. The result is a much more refined profile of psychopathology than could be achieved with a categorical approach. For etiological research this may improve the chances to find a relationship between for instance biological factors and psychopathology. For clinical practice it allows (provided enough data are available) more refined choices for treatment and a better prediction of the prognosis.

The aim of this thesis was to investigate diagnostic heterogeneity and to test the feasibility of dimensional models in a large, real-life group of psychiatric outpatients with mood, anxiety and / or somatoform disorders and to develop a dimensional model that overcomes the disadvantages of existing ones. Before discussing the results, the major findings will be summarized.

7.2 Summary of major findings

As the data of all patient samples in this thesis were collected with Routine Outcome Monitoring (ROM), we first described this method in detail in Chapter 2. Although they initially had their reservations, most therapists considered ROM to be an important adjunct to diagnostics and treatment outcome evaluation. In addition, ROM furthers research as the data can be used to study

the phenomenology of psychiatric disorders and the outcome of treatments delivered in everyday practice. Implementation of ROM in outpatients with depressive, anxiety and somatoform disorders therefore seems to be feasible and useful.

Next, we investigated whether in this patient sample a high rate of comorbidity (as discussed in 7.1) could indeed be found. We analysed the prevalence of axis 1 DSM-IV disorders in a group of 3798 outpatients who had had ROM-assessments. According to the MINI-Plus (part of ROM), 1,618 patients (42.6%) met criteria for a single mood, anxiety, or somatoform (MAS) disorder, but nearly the same number, 1,556 patients (41.0%), had more than one concurrent MAS disorder: 967 patients (25.5%) had two comorbid disorders, 403 patients (10.6%) had three, and 186 patients (4.9%) had four or more. This high prevalence of different types of comorbidity signifies heterogeneity.

In Chapter 3 we examined whether the comorbidity discussed in Chapter 2 is merely the coming together of two or more disorders in the same patient or whether the whole is more than the sum of its parts. To do so, we compared the scores of patient groups defined by the categorical diagnoses on several severity assessments, This approach is 'semi-dimensional' as it stays close to the diagnostic categories but allows quantification. We found that depression severity in the comorbid group was higher than in the pure depression group and that anxiety severity in the comorbid group was higher than in the pure anxiety group. This study also revealed that the mean scores on the anxiety measures did not differ significantly between patients with a pure depression and patients with a pure anxiety disorder. These results show that, with respect to symptom severity, comorbidity is more than simply the sum of the disorders.

We also wanted to go beyond categorical diagnoses and explore a more fully dimensional model with dimensions not necessarily coupled to the diagnostic categories of depressive and anxiety disorders. We chose an already existing model as point of departure: the tripartite model of Watson and Clark. This model proposes that there is one nonspecific general distress factor (negative affect), common to both mood and anxiety disorders, and two additional factors specific to anxiety disorders and depression. The three dimensions of the tripartite model can be measured with the MASQ (Mood and Anxiety Symptom Questionnaire). In order to do research on the tripartite model in Dutch samples, a translation of the MASQ was needed. In Chapter 4 the Dutch adaptation of the MASQ is presented and the applicability of the tripartite model on our sample is tested. 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.

Although the tripartite model has inspired a large body of research, it has met some criticism as well. A major point of critique is that depression is well

covered with lack of positive affect and negative affect (a nonspecific aspect of the disorder). However, the same cannot be said for anxiety, as the dimension 'somatic arousal' that is specific for anxiety does not cover all anxiety disorders but mainly covers panic disorder. Somatic arousal is too narrow as a conceptualization of anxiety, ignoring other important aspects of anxiety such as anxious apprehension, worry, phobic anxiety and/or avoidance.

In Chapter 5, we present a first model that contains clearly distinguishable constructs, and includes main aspects of common mental disorders in outpatients. Our aim was to cover anxiety more adequately than the tripartite model does. We used items of the Mood and Anxiety Symptom Questionnaire (Watson & Clark, 1991) and items of the Brief Symptom Inventory (Derogatis, 1975). A model with five dimensions was found: depressed mood, lack of positive affect, somatic arousal, phobic fear and hostility. The validity of the model was supported by the following findings: The scales appeared capable to differentiate between patients with either a mood or an anxiety disorder. Low positive affect and phobic fear were the best discriminators between depressed patients and patients with an anxiety disorder. Within the anxiety disorders, somatic arousal was specific for patients with panic disorder. Phobic fear was associated with panic disorder, simple phobia and social anxiety disorder, but not with generalized anxiety disorder.

Whereas the study described in Chapter 5 validated the model by comparing the dimensions to the categories of the DSM-IV (disorder-based approach), in Chapter 6 we took a step further away from the DSM-IV and closer towards a 'true dimensional model' (symptom-based approach). 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 anxiety subscales of the Brief Symptom Inventory (Derogatis, 1975), to measure fearfulness and (3.) newly designed items to measure anxious apprehension. By using two different patient samples to develop and evaluate this second model, we arrived at a 6-factor model: feelings of worthlessness, fatigue, somatic arousal, anxious apprehension, phobic fear and tension. Somatic arousal, anxious apprehension and phobic fear are all clearly anxiety-like constructs. Thus, instead of only the single anxiety dimension of the tripartite model (somatic arousal), the present model distinguishes three groups of symptoms. Each individual factor and the total of factors can be regarded as unidimensional measurement scales, and this model can describe the clinical state of patients more specifically than the tripartite model.

7.3 Discussion

7.3.1 The results

The studies presented in this thesis are explorations along the road to a fully dimensional model of psychopathology. Below we will discuss what we contributed and what our contributions mean.

First of all, we showed in our own sample of secondary care outpatients with depression, anxiety and somatoform disorders that comorbidity is highly prevalent. This is in line with the findings of an extensive body of literature on comorbidity. High rates of comorbidity between anxiety disorders and depression have also been reported in the general population (Kessler et al., 1996), in primary care (Roca et al., 2009) and in secondary care (Brown et al., 2001). In fact, comorbidity of depressive and anxiety disorders is so prevalent that it is no coincidence (Kessler et al., 1996). These findings launched research into the existence of psychopathological dimensions common to anxiety and depression (Clark & Watson, 1991) and into genetic overlap (Kendler, 1996). The psychopathological dimensions common to depression and anxiety disorders will be discussed later, after the role of the dimensional approach in assessing the severity of comorbid disorders. The possible genetic overlap goes beyond the scope of this paper and will not be discussed further.

We showed that in depression and anxiety disorders comorbidity is more than simply the sum of diagnoses. For instance, we found that some symptoms of comorbid occurring disorders, are more severe than if the disorders occur alone. This has also been reported in other studies (e.g. Dalrymple & Zimmerman, 2007; Fava et al., 2004; Kaufman & Charney, 2000) but has never been studied for comorbid depression and anxiety disorders and single depression and single anxiety disorders in the same clinical sample. Together with the symptom heterogeneity possible in patients with the same diagnosis (see 7.1), the results suggest that the categorical diagnoses as defined in the DSM-IV are too indistinct. Assessing symptom severity may make etiological research more fruitful and may also help to find more effective treatments. We will discuss this more comprehensively further on. For now, it suffices to say that the proposal to include the assessment of symptom severity in the DSM-5 signifies growing support for this view.

What is the best instrument to assess dimensions of depressive and anxiety symptoms? In fact, every multi-item questionnaire on depressive and / or anxiety symptoms yields a quantitative assessment of one or more aspects of psychopathology and thus may qualify for the assessment of these dimensions. We chose the Mood- and Anxiety Symptom Questionnaire (MASQ) of Watson and Clark (1991) as it is based on their tripartite model. The significance of this model lies in the fact that it tries to take into account the overlap as well as the diversity in psychopathology in subjects suffering from depressive and / or anxiety disorders. The overlap is assessed with a non-specific distress factor

(negative effect) and the diversity with a factor specific for depression and one for anxiety. The MASQ also was chosen because of its use in many studies, also with respect to etiology, not only with adult patients (Marshall et al., 2003; Joiner et al., 1999; Keogh & Reidy, 2000) but also in child psychiatry populations (Chorpita & Daleiden, 2002).

As we made the MASQ the central assessment tool of our further investigations, it was important to have at our disposal a translated and psychometrically sound Dutch version. We carefully translated the MASQ and demonstrated good reliability and validity of this Dutch version in a large sample of 950 outpatients referred to secondary care because of mood, anxiety and/or somatoform (MAS-) disorders and 200 respondents from the general population. We did not include inpatients, primary care patients and patients with other disorders as our research focused on outpatients with MAS-disorders. However, in the future the MASQ should also be evaluated psychometrically in those groups of subjects. The present analysis showed that the factor structure of the MASQ with three factors was preserved in the Dutch translation. Factor-loadings of items and allocation of items to subscales was similar to results of Watson and Clark with US clinical samples and with patient samples from Great Britain (Keogh & Reidy, 2000). Recently, our group has developed and evaluated a shortened 30-item version, called the MASQ-D30, thereby increasing the feasibility of its incorporation in an assessment battery for ROM (Wardenaar et al., 2010).

Translation of the MASQ was not the primary aim of our study, but rather a means to an end. Our main aim was to remediate the shortcomings of the tripartite model and the MASQ. The original authors recommended already in 1998 to view in future research "individual disorders as representing unique *combinations* of different types of symptoms, with each type showing varying degrees of nonspecificity and with no type being entirely unique to any single disorder" (Mineka et al., 1998, p.398). We operationalized this idea, by developing symptom scales that include the more unique symptoms of specific mood and anxiety disorders in addition to common symptom scales. As described in chapter 5, by adding items of the BSI to the MASQ the new questionnaire was able to distinguish three groups of symptoms, each one specific to a different kind of anxiety disorder (panic disorder, GAD, and phobic disorders) instead of only the single anxiety dimension of the tripartite model (somatic arousal).

However, remediating the shortcomings of the MASQ in differentiating between the various DSM-IV categories of mood and anxiety disorders was not our final goal either. Rather, we set out to develop a broad dimensional model, not taking DSM-IV diagnoses as a point of departure nor taking specificity to particular DSM-IV diagnoses as the best sign of validity. We aimed for a multidimensional model to characterize individual patients in

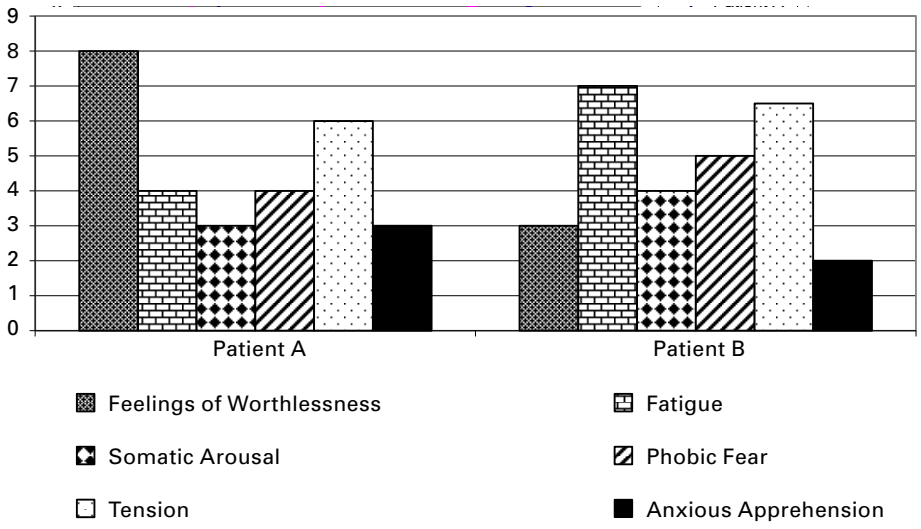


Figure 7.1 Fictional example of a 'symptom profile' for two patients, both diagnosed with Major Depressive Disorder according to the DSM-IV.

terms of their specific symptom profile by including extra symptoms to a self-report instrument in order to cover additional dimensions. As described in chapter 6, this resulted in a 6-factor model: feelings of worthlessness, fatigue, somatic arousal, anxious apprehension, phobic fear and tension. This model reveals differences in symptom profiles between patients who, according to the DSM-IV would all have been diagnosed with MDD. Figure 7.1 gives a graphic representation of the symptom profile of the two exemplary patients we introduced in the introduction of this thesis. Not only patients with MDD, but also patients with anxiety disorders and patients with comorbid depression and anxiety, can be characterized with the same 6 factors.

The dimensions may be a fruitful basis for future research into prognostic factors of treatment response. It may well be that an optimal match exists between symptom profiles and treatment modality. By assessing large groups of patients before and after treatment with, for instance, selective serotonin reuptake inhibitors (SSRIs), it will become possible to determine which profile(s) are most sensitive for these drugs. Finding the most appropriate treatment (pharmacologically or psychotherapeutic) can be a lengthy trial-and-error process. Matching of patient characteristics to treatments is the next step in improving evidence based medicine in psychiatry (Beutler, Forrester, Gallagher-Thomson, Thompson, & Tomlins, 2012). Eventually, this information

will help to address the famous question first raised by Gordon Paul: "what treatment, by whom, is most effective for this individual with that specific problem, and under which set of circumstances?" (Paul, 1967, p.111).

The model presented in Chapter 6 is a hierarchical model with a bifactor structure (see for a graphic representation: model 5 in Figure 6.1). Confirmatory factor analyses showed that this bifactor hierarchical model with a general severity factor and six specific factors fitted best to the data, compared to several other models. 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. With a hierarchical common factor it is possible to determine severity and to differentiate between non-patients and patients. The dimensions can be used to form a unique symptom profile for each patient to differentiate within patients.

The tripartite model is a unifactorial model with 3 factors. In a unifactorial model a set of items load on one factor (see for examples of unifactorial models: model 1, 2 and 3 in Figure 6.1). An important feature of the tripartite model (Clark & Watson, 1991) is the difference that is made between general and specific distress. The factor negative affect in the model is presented as general distress, whereas 'lack of positive affect' and 'somatic arousal' are presented as dimensions of distress specific for respectively depression and anxiety disorders.

This difference in general and specific distress is also an element in several other models which were previously suggested as reaction to the critiques on the tripartite model (Brown et al., 1998; Mineka et al., 1998; Zinbarg & Barlow, 1996). There is consensus that both general and specific components are needed to fully represent the variation observed among mood and anxiety disorders (Simms et al., 2008). It is confusing, however, that the terms 'general' and 'specific' are not used consistently in literature. They can be used to refer to the content of a factor, and also to refer to their place in a higher-order or hierarchical model. For example, NA in the tripartite model is not a general factor according to the methodological structure of the model (a unifactorial model does not contain any general, higher order or common factor). However the content of NA is general (since it represents general distress and not a specific symptom), and in this context the term *specific* means specific to a (group of) disorder(s) or patients. We advocate to refer to NA as non-specific, and to avoid the term general in this context. In a hierarchical, bifactor model, we prefer to speak of unique and common factors.

We concluded 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 unique factors in our model describe a patients' 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 the symptomatology of depression and anxiety has a hierarchical structure.

7.3.2 Limitations

The results should be interpreted in the light of some limitations. First, we took the symptoms mentioned in the DSM definitions of mood and anxiety disorders as a starting point. In theory, other symptom dimensions could be of importance to describe the phenotype of mental disfunction which show different associations with etiology, the course and treatment of mental problems. However, depressivity and anxiety are universal notions that are elaborated by psychology and psychiatry. Positive and negative affect, the two dimensions from the circumplex model of affect (Watson & Tellegen, 1985) that Watson and Clark used in their tripartite model (Clark & Watson, 1991), are also strongly embedded and universal (Russell & Lewicka, 1989).

Second, the described studies are limited to the common mental disorders depression and anxiety disorders (both internalizing disorders). Although the difference between internalizing and externalizing disorders is often confirmed in research (Kessler et al., 2005; Kotov et al., 2011; Vollebergh et al., 2001), symptoms of externalizing disorders are also present in patients with internalizing disorders (and vice versa). For example, Koh and colleagues (2002) found a predominance of anger in depressive disorders compared with anxiety disorders and somatoform disorders. Therefore, it is worthwhile to extend the model with externalizing dimensions (Krueger et al., 2005) comprising concepts such as 'anger' or 'aggression' (Pasquini et al., 2004; Picardi et al., 2004). We made a start with the dimension 'hostility' in the model presented in Chapter 5.

Third, the results only apply to outpatients before the start of their treatment. This implies that the findings cannot be generalized to inpatients or to persons with "normal" or nonpathological levels of anxiety and depression as a general model of affect. To make a model that can be generalized to the normal population it might be useful to measure all dimensions with both negatively and positively formulated items. After all, the measurement range of the dimensions will be wider when positively formulated items are included as well.

Fourth, the results only apply to patients with specific demographic characteristics. Approximately 80 percent of the patients in our samples were born in The Netherlands, as were both their parents. On top of that, a condition to participate in ROM was to master the Dutch language well, both spoken en written. Therefore, no statements can be made about to what extent our results apply to patients with different ethnic backgrounds or literacy. Moreover, the results cannot be generalized to children and elderly, since we used patient samples of adults only.

Finally, this thesis focused on the reliability and the internal validity of the multidimensional models, not on the external validity. We did not investigate to what extent the dimensions correlate with biological factors like cortisol levels or polymorphisms and to what extent they predict treatment success and the course of the mental problems. This was done, but as yet with the original dimensions of the MASQ, by other members of the research group of the LUMC department of psychiatry. Van Veen et al. (2013) found that childhood traumas have different effects on the MASQ dimensions, whereas most adult life events are associated with all three. Wardenaar et al. (2012) showed that MASQ dimensions predicted the future 2-year course of depression and anxiety. Importantly, the dimensions yield predictive information on top of DSM-IV diagnoses. Luppino et al. (2011) demonstrated a strong association of most components of the metabolic syndrome with the SA dimension, but not the PA and NA dimension of the MASQ. Veen et al. (2011) and Wardenaar et al (2011) both found non-linear relations between the cortisol awakening rise (CAR) and dimensions of the MASQ, which could explain previous inconsistent findings regarding HPA-axis activity in depressed patients. And last but not least, Veen et al. (2012) showed that MASQ-dimensions were each associated with specific gene sets. It can be concluded that the external validity of the original MASQ dimensions is promising. It will be interesting to investigate the external validity of the extended dimensional model presented in this thesis.

7.3.3 Future perspectives

Before adopting a dimensional approach on a large scale, the superiority of the dimensional approach to the DSM-IV for the characterization of patients, the investigation of the etiology and the clinical utility needs to be demonstrated (First, 2005). Future research has to show whether a dimensional profile is indeed useful in deciding what the main target for treatment should be and what kind of treatment is indicated. For example, an overactive sympathic nervous system as revealed by high anxious arousal may require a different pharmacotherapeutic approach, while a high propensity to worry may suggest psychosocial therapy. With the original scales of the tripartite model, the first progress in using dimensions in research into etiology is already made. For example, Wardenaar found nonlinear associations between characteristics of the stress-system (cortisol awakening curve) and the dimensions of the tripartite model (2011) in a sample of outpatients.

In most current research into the etiology of common mental disorders, patients are compared to controls regarding the presence of specific genes, or other biological or psychological variables. Most commonly, this is done with a categorical "mindset": the presence or absence of a trait or biological marker is investigated in persons with or without a diagnosis (e.g., patients with a major depressive disorder as compared to controls). A dimensional model

however requires a correlational "mind set": etiological factors (themselves often measured on a continuous scale) may be strongly correlated with some (combinations of) dimensions and less strongly with others, irrespective of the categorical diagnoses the patients have. Thus, in research aimed for instance at the endophenotypes of psychopathology we might find the "anxious apprehension profile" rather than "anxiety disorder". It is quite a challenge for researchers to switch from a categorical to a correlational mind set. After all, all humans have a strong tendency toward categorization as we are more inclined to separate and sort things (safe - unsafe, edible - not edible) so we know how to navigate in the world around us. This 'mental categorization' is one of the first stages in our cognitive development, and starts at a very young age (Piaget, 1962). Besides our early learned custom to think in categories, another difficulty is that dimensions are much more complex to depict than categories, especially when more than 3 dimensions are involved.

A pleasant consequence of using multi dimensional models in research is that statistical power is usually substantially enhanced if true variance in affect scores is assessed and preserved in the analysis. This is easily demonstrated by comparing a dimensional depression score with the dichotomous categorization of depressed vs. non-depressed. With the latter, much information is sacrificed which would have been preserved in the former.

On the other hand, using dimensions in research means that the required statistical methods are more complex than when using categories. A t-test between two groups (e.g. not depressed vs. depressed) to test for a significant difference on another variable won't do. Multiple regression analyses with special attention for interaction effects will be necessary. Fortunately, these techniques are now available in statistical software.

It is preferable that all symptom dimensions are analysed simultaneously. Although the correlations between the dimensions are relatively low, they are correlated. If one would study or analyse them one by one, correction for this correlation is not possible and the wrong conclusions might be drawn. For example the conclusion might be drawn that a specific treatment does not have an effect on both dimensions A and B separately, while the interaction effect is missed.

A more general point about the use of dimensions in research is that they break with the simple tradition of comparing etiological factors between healthy and diseased groups. This can lead to a disruption in research efforts; e.g. combined meta-analyses are not possible on studies that use DSM-categories and studies that use a dimensional approach. This drawback can be prevented by a combined approach (using both the DSM and dimensional measures). For research, a combined approach has great benefits over using the categorical system solely (Brown & Barlow, 2005). It is already seen more and more in research that although the selection of patients for a research group

is still determined by the DSM-criteria, dimensional measures are added to the research design. An important advantage of this development is that individual differences between patients within one DSM-category are acknowledged (e.g. in analyses of experimental findings the scores on the dimensional measures can be taken into account as covariates). As Kaufman and Charney concluded: 'the use of categorical diagnostic approaches and dimensional rating scale in tandem will facilitate identification of meaningful phenotypes for future genetic, biochemical, neuroimaging, and treatment studies' (pag.73, Kaufman & Charney, 2000).

For the successor of the DSM-IV, the DSM-5 (published in May, 2013) it was suggested to combine categories with dimensional measures. The DSM-5 Work Groups were considering an additional way to help the clinician capture the symptoms and severity of mental illnesses, by using dimensional assessments. These would allow clinicians to systematically evaluate patients on the full range of symptoms they may be experiencing. For instance, information about depressed mood, anxiety level, quality of sleep, and substance use would be important for clinicians to know regardless of the patient's diagnosis. Dimensional assessments would allow clinicians to rate both the presence and the severity of the symptoms, such as "very severe," "severe," "moderate", "mild", or "absent". It would encourage mental health professionals to document all of a patient's symptoms and not just those that were tied to their primary diagnosis.

Adopting dimensions in the DSM-5 holds much promise. It is a start to advocate dimensions in the field. And, although clinical utility in the sense that assessed dimensions can be used to decide which treatment is most effective for a specific patient is not available yet, there is however already a benefit for the clinician in assessing dimensional measures next to the DSM-categories. When using both a diagnostic interview and several dimensional (severity) measures at intake, the clinician gets useful insights into the symptoms profile of each patient at intake and at follow-up (ROM), and therewith into the effect of the chosen treatment on different sorts of symptoms for the patient at hand.

We suggest using the same dimensional measures within all DSM-categories of common mental disorders. For example, only when anxiety is measured dimensionally in both patients with a depression as well as those with an anxiety disorder, analyses can be done in all categories simultaneously. Only then, research can be done without the restrictions of the DSM. ROM as implemented in 2002 in Leiden, proved to be a very useful instrument to measure dimensions and categories, and combined with biological data enhance our insight in the complex relationship between depression and anxiety and their common and distinctive etiological factors.

The multi-dimensional models presented in this thesis are limited to symptoms that were present at the time of the assessment. The questionnaires

used, ask the patient to report the level of presence of each symptom in the week prior to the assessment. A symptom profile generated with the models therefore does not contain any information about the history of the patient (duration, recurrence, familiarity etcetera). Determining what phase of the clinical course of the disorder a patients is in (staging) is very important. A clinical staging model, already widely used in oncology, could improve the utility of diagnostic characterisation in psychiatry as well, with emerging disorders (McGorry et al., 2007). Staging models are based on the fact that response to treatment is generally better when it is introduced early in the course of the illness. It assumes that earlier stages have better prognosis and require simpler therapeutic regimens (Vieta, Reinares, & Rosa, 2011). It would be ideal to include symptom-profiles ('profiling') in 'staging'. Routine Outcome Monitoring is an important instrument for developing staging and profiling in psychiatry (Zitman, 2012).

We believe that the main focus for the next years should be on research on profiling and staging with the aim to determine those factors that predict the evolution of symptoms and the effective treatment. This kind of research has the best chance of being successful when various research groups cooperate and find consensus about research designs and variables used and how to conceptualize them optimally. A dimensional approach to psychopathology is expected to be more successful than the traditional categorical approach, as it is a far better representation of the richness of clinical phenomena.

REFERENCE LIST

Aaronson, N.K., Muller, M., Cohen, P.D., Essink-Bot, M.L., Fekkes, M., Sanderman, R. et al. (1998). Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *Journal of Clinical Epidemiology, 51,* 1055-1068.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders (2th ed.)*. Washington, DC: American Psychiatric Publishing Inc.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders (3th ed. revised, IIIR)*. Washington, DC: American Psychiatric Publishing Inc.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, DC: American Psychiatric Publishing Inc.

Andrews, G. (1996). Comorbidity and the general neurotic syndrome. *British Journal of Psychiatry,* 76-84.

Arrindel, W. & Ettema, J. (1986). *SCL-90: Handleiding bij een multidimensionele psychopathologie-indicator. (SCL-90: Manual for a multidimensional indicator of psychopathology).* Lisse. The Netherlands: Swets & Zeitlinger.

Barkham, M., Culverwell, A., Spindler, K., & Twigg, E. (2005). The CORE-OM in an older adult population: psychometric status, acceptability, and feasibility. *Aging Mental Health, 9,* 235-245.

Barlow, D.H., Gorman, J.M., Shear, M.K., & Woods, S.W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial. *Journal of the American Medical Association, 283,* 2529-2536.

Beck, A.T. & Steer, R.A. (1987). *Manual for the revised Beck Depression Inventory*. San Antonio, TX: Psychological Corporation.

Bedford, A. (1997). On Clark-Watson's tripartite model of anxiety and depression. *Psychological Reports, 80,* 125-126.

Belzer, K. & Schneier, F.R. (2004). Comorbidity of anxiety and depressive disorders: issues in conceptualization, assessment, and treatment. *Journal of Psychiatric Practice, 10,* 296-306.

Bentler, P.M. (2006). *EQS 6 Structural Equations Program Manuel*. Multivariate Software Inc, Encino, CA.

Beuke, C.J., Fischer, R., & McDowall, J. (2003). Anxiety and depression: Why and how to measure their separate effects. *Clinical Psychology Review, 23,* 831-848.

Beutler, L.E., Forrester, B., Gallagher-Thomson, D., Thompson, L., & Tomlins, J.B. (2012). Common, specific, and treatment fit variables in psychotherapy outcome. *Journal of Psychotherapy Integration, 22,* 255-281.

Boschen, M.J. & Oei, T.P. (2006). Factor structure of the Mood and Anxiety Symptom Questionnaire does not generalize to an anxious/depressed sample. *Australian and New Zealand Journal of Psychiatry, 40,* 1016-1024.

Boschen, M.J. & Oei, T.P. (2007). Discriminant validity of the MASQ in a clinical sample. *Psychiatric Research, 150,* 163-171.

Brown, T.A., Campbell, L.A., Lehman, C.L., Grisham, J.R., & Mancill, R.B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110,* 585-599.

Brown, T.A. & Leyfer, O. (2009). Classification of Anxiety Disorders. In D.J.Stein, E. Hollander, & B.O. Rothbaum (Eds.), *Textbook of Anxiety Disorders* (2nd ed., pp. 17-35). American Psychiatric Publishing, Inc.

Brown, T.A. & Barlow, D.H. (2005). Dimensional Versus Categorical Classification of Mental Disorders in the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders and Beyond: Comment on the Special Section. *Journal of Abnormal Psychology, 114,* 551-556.

Brown, T.A., Chorpita, B.F., & Barlow, D.H. (1998). Structural Relationships Among Dimensions of the DSM-IV Anxiety and Mood Disorders and Dimensions of Negative Affect, Positive Affect, and Autonomic Arousal. *Journal of Abnormal Psychology, 107,* 179-192.

Buckby, J.A., Cotton, S.M., Cosgrave, E.M., Killackey, E.J., & Yung, A.R. (2008). A factor analytic investigation of the Tripartite model of affect in a clinical sample of young Australians. *BMC Psychiatry, 8: 79*.

Buckby, J.A., Yung, A.R., Cosgrave, E.M., & Cotton, S.M. (2007). Distinguishing between anxiety and depression using the Mood and Anxiety Symptom Questionnaire (MASQ). *British Journal of Clinical Psychology, 46,* 235-239.

Burgess, P.M., Pirkis, J.E., Slade, T.N., Johnston, A.K., Meadows, G.N., & Gunn, J.M. (2009). Service use for mental health problems: findings from the 2007 National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry, 43,* 615-623.

Burns, D.D. & Eidelson, R.J. (1998). Why are depression and anxiety correlated? A test of the tripartite model. *Journal of Consulting and Clinical Psychology, 66,* 461-473.

Byrne, B.A. (2006). *Structural Equation Modeling with EQS: Basic Concepts, Applications and Programming.* New Jersey: Lawrence Erlbaum Associates.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Catell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245-276.

Chorpita, B.F. (2002). The tripartite model and dimensions of anxiety and depression: an examination of structure in a large school sample. *Journal of Abnormal Child Psychology, 30,* 177-190.

Chorpita, B.F. & Daleiden, E.L. (2002). Tripartite dimensions of emotion in a child clinical sample: measurement strategies and implications for clinical utility. *Journal of Consulting and Clinical Psychology, 70,* 1150-1160.

Clark, D.M., Fairburn, C.G., & Wessely, S. (2008). Psychological treatment outcomes in routine NHS services: a commentary on Stiles et al. (2007). *Psychological Medicine, 38,* 629-634.

Clark, L.A. & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336.

Clark, L.A. (2005). Temperament as a Unifying Basis for Personality and Psychopathology. *Journal of Abnormal Psychology, 114,* 505-521.

Cuthbert, B.N. (2005). Dimensional Models of Psychopathology: Research Agenda and Clinical Utility. *Journal of Abnormal Psychology, 114,* 565-569.

Dalrymple, K.L. & Zimmerman, M. (2007). Does comorbid Social Anxiety Disorder impact the clinical presentation of principal Major Depressive Disorder? *Journal of Affective Disorders, 100,* 241-247.

De Beurs E., Rinne, T., van, K.D., Verheul, R., & Andrea, H. (2009). Reliability and validity of the Dutch Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF), a shortened version of the DAPP-Basic Questionnaire. *Journal of Personality Disorders, 23,* 308-326.

De Beurs, E. (2005). *De Brief Symptom Inventory; Handleiding (The Brief Symptom Inventory; Manual)*. Leiden: Pits Publishers.

De Beurs, E., Beekman, A.T., van Balkom, A.J., Deeg, D.J., Van Dyck, R., & van Tilburg, W. (1999). Consequences of anxiety in older persons: its effect on disability, well-being and use of health services. *Psychological Medicine, 29,* 583-593.

De Beurs, E., Comijs, H., Twisk, J.W., Sonnenberg, C., Beekman, A.T., & Deeg, D. (2005). Stability and change of emotional functioning in late life: modelling of vulnerability profiles. *Journal of Affective Disorders, 84,* 53-62.

De Beurs, E., Den Hollander-Gijsman, M.E., Helmich, S., & Zitman, F.G. (2007). 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, 45,* 1609-1617.

De Beurs, E., Den Hollander-Gijsman, M.E., Van Rood, Y.R., Van der Wee, N.J.A., Giltay, E.J., Van Noorden, M.S. et al. (2011). Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology and Psychotherapy, 18,* 1-12.

De Graaf, R., Bijl, R.V., Smit, F., Vollebergh, W.A., & Spijker, J. (2002). Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: findings from the Netherlands Mental Health Survey and Incidence Study. *American Journal of Psychiatry, 159,* 620-629.

Den Hollander-Gijsman, M.E., De Beurs, E., Van der Wee, N.J.A., Van Rood, Y.R., & Zitman, F. G. (2010). Distinguishing between depression and anxiety: a proposal for an extension of the tripartite model. *European Psychiatry, 25,* 197-205.

Derogatis, L.R. (1975). *The Brief Symptom Inventory*. Baltimore, MD.: Clinical Psychometric Research.

Derogatis, L.R., Lipman, R.S., & Covi, L. (1973). SCL-90: an outpatient psychiatric rating scale-preliminary report. *Psychopharmacology Bulletin, 9,* 13-28.

Derogatis, L.R. & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine, 13,* 595-605.

Dillmann, D.A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.

Ellwood, P.M. (1988). Outcomes management. A technology of patient experience. *New England Journal of Medicine, 318,* 1549-1556.

Emmanuel, J., Simmonds, S., & Tyrer, P. (1998). Systematic review of the outcome of anxiety and depressive disorders. *British Journal of Psychiatry, 173 (suppl 34),* 35-41.

Endicott, J., Spitzer, R.L., Fleiss, J.L., & Cohen, J. (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33,* 766-771.

Fava, M., Alpert, J.E., Carmin, C.N., Wisniewski, S.R., Trivedi, M.H., Biggs, M.M. et al. (2004). Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D. *Psychological Medicine, 34,* 1299-1308.

First, M.B. (2005). Clinical Utility: A Prerequisite for the Adoption of a Dimensional Approach in DSM. *Journal of Abnormal Psychology, 114,* 560-564.

Floyd, F.J. & Widaman, K.F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment, 7,* 286-299.

Gilbody, S.M., House, A.O., & Sheldon, T.A. (2002). Outcomes research in mental health - Systematic review. *British Journal of Psychiatry, 181,* 8-16.

Goekoop, J.G., Knopper-Van der Klein, E.A., Hoeksema, T., Klinkhamer, R.A., Van Gaalen, H.A., & van der Velde, E.A. (1991). The interrater reliability of a Dutch version of the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica, 83,* 202-205.

Goekoop, J.G., De Beurs, E., & Zitman, F.G. (2007). Four-dimensional structure underlying scales for depression anxiety and retardation: emergence of trapped anger and scale improvements. *Comprehensive Psychiatry, 48,* 192-198.

Goldberg, D. (2000). Plato versus Aristotle: categorical and dimensional models for common mental disorders. *Comprehensive Psychiatry, 41,* 8-13.

Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology* Rockville, MD.: National Institute of Mental Health.

Heller, W., Nitschke, J.B., Etienne, M.A., & Miller, G.A. (1997). Patterns of regional brain activity differentiate types of anxiety. *Journal of Abnormal Psychology, 106,* 376-385.

Holloway, F. (2002). Outcome measurement in mental health - welcome to the revolution. *British Journal of Psychiatry, 181,* 1-2.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Jacobson, N.S., Roberts, L.J., Berns, S.B., & McGlinchey, J.B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67,* 300-307.

Jensen-Doss, A. & Weisz, J.R. (2008). Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics. *Journal of Consulting and Clinical Psychology, 76,* 711-722.

Joffe, R.T., Bagby, R.M., & Levitt, A. (1993). Anxious and nonanxious depression. *American Journal of Psychiatry, 150,* 1257-1258.

Joiner, T.E. & Lonigan, C.J. (2000). Tripartite model of depression and anxiety in youth psychiatric inpatients: relations with diagnostic status and future symptoms. *Journal of Clinical Child Psychology, 29,* 372-382.

Joiner, T.E., Steer, R.A., Beck, A.T., Schmidt, N.B., Rudd, M.D., & Catanzaro, S.J. (1999). Physiological hyperarousal: construct validity of a central aspect of the tripartite model of depression and anxiety. *Journal of Abnormal Psychology, 108,* 290-298.

Kaufman, J. & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety, 12 Suppl 1,* 69-76.

Kendler, K.S. (1996). Major depression and generalised anxiety disorder. Same genes, (partly)different environments--revisited. *British Journal of Psychiatry.Supplement,* 68-75.

Keogh, E. & Reidy, J. (2000). Exploring the factor structure of the Mood and Anxiety Symptom Questionnaire (MASQ). *Journal of Personality Assessment, 74,* 106-125.

Kessler, R.C., Barker, P.R., Colpe, L.J., Epstein, J.F., Gfroerer, J.C., Hiripi, E. et al. (2003). Screening for serious mental illness in the general population. *Arch.Gen.Psychiatry, 60,* 184-189.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 593-602.

Kessler, R.C., Berglund, P.A., Dewit, D.J., Ustun, T.B., Wang, P.S., & Wittchen, H.U. (2002). Distinguishing generalized anxiety disorder from major depression: prevalence and impairment from current pure and comorbid disorders in the US and Ontario. *International Journal of Methods in Psychiatric Research, 11,* 99-111.

Kessler, R.C., Nelson, C.B., McGonagle, K.A., Liu, J., Swartz, M., & Blazer, D.G. (1996). Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. *British Journal of Psychiatry, 168,* 17-30.

Kessler, R.C., Chiu, W.T., Demler, O., & Walters, E.E. (2005). Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 617-627.

Knaup, C.M., Koesters, D., Schoefer, T., Becker, T., & Pushner, B. (2009). Effect of feedback on treatment outcome in specialist mental health care: meta-analysis. *British Journal of Psychiatry, 195,* 15-22.

Koh, K.B., Kim, C.H., & Park, J.K. (2002). Predominance of anger in depressive disorders compared with anxiety disorders and somatoform disorders. *Journal of Clinical Psychiatry, 63,* 486-492.

Kotov, R., Ruggero, C.J., Krueger, R.F., Watson, D., Yuan, Q., & Zimmerman, M. (2011). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry, 68,* 1003-1011.

Krueger, R.F. (1999). The structure of common mental disorders. *Archives of General Psychiatry, 56,* 921-926.

Krueger, R.F., Markon, K.E., Patrick, C.J., & Iacono, W.G. (2005). Externalizing Psychopathology in Adulthood: A Dimensional-Spectrum Conceptualization and Its Implications for DSM-V. *Journal of Abnormal Psychology, 114,* 537-550.

Krueger, R.F., Watson, D., & Barlow, D.H. (2005). Introduction to the Special Section: Toward a Dimensionally Based Taxonomy of Psychopathology. *Journal of Abnormal Psychology, 114,* 491-493.

Kupfer, D.J.M. (2005). Dimensional Models for Research and Diagnosis: A Current Dilemma. *Journal of Abnormal Psychology, 114,* 557-559.

Lambert, M.J., Burlingame, G.M., Umphress, V.J., Hanssen, N., Vermeersch, D., Clause, G. et al. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 3,* 249-258.

Lambert, M.J., Hansen, N.B., & Finch, A.E. (2001). Patient-focused research: using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69,* 159-172.

Lambert, M.J., Harmon, C., Slade, K., Whipple, J.L., & Hawkins, E.J. (2005). Providing feedback to psychotherapists on their patients' progress: clinical results and practice suggestions. *Journal of Clinical Psychology, 61,* 165-174.

Lecrubier, Y., Sheehan, D.V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K.H. et al. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry, 12,* 224-231.

Lee, W., Jones, L., Goodman, R., & Heyman, I. (2005). Broad outcome measures may underestimate effectiveness: An instrument comparison study. *Child and Adolescent Mental Health, 10,* 143-144.

Livesley, W.J. & Jackson, D.N. (2006). *Manual for the Dimensional Assessemnt of Personality Problems - basic questionnaire*. Port Huron, Michigan: Sigma.

Luppino, F.S., van Reedt Dortland, A.K., Wardenaar, K.J., Bouvy, P.F., Giltay, E.J., Zitman, F.G. et al. (2011). Symptom dimensions of depression and anxiety and the metabolic syndrome. *Psychosomatic Medicine, 73,* 257-264.

MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19-40.

Marshall, G.N., Sherbourne, C.D., Meredith, L.S., Camp, P., & Hays, R.D. (2003). The tripartite model of anxiety and depression: symptom structure in depressive and hypertensive patient groups. *Journal of Personality Assessment, 80,* 139-153.

McGorry, P.D., Purcell, R., Hickie, I.B., Yung, A.R., Pantelis, C., & Jackson, H.J. (2007). Clinical staging: a heuristic model for psychiatry and youth mental health. *Medical Journal of Australia, 187,* S40-S42.

Miller, S.D., Duncan, B.L., Sorrell, R., & Brown, G.S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61,* 199-208.

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

Montgomery, S.A. & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry, 134,* 382-389.

Moser, C.A. & Kalton, G. (1979). *Survey methods in social investigation*. London: Heinemann.

Nitschke, J.B., Heller, W., Imig J.C., McDonald P., & Miller, G. A. (2001). Distinquishing Dimensions of Anxiety and Depression. *Cognitive Therapy and Research, 25,* 1-22.

Nitschke, J.B., Heller, W., Palmieri, P.A., & Miller, G.A. (1999). Contrasting patterns of brain activity in anxious apprehension and anxious arousal. *Psychophysiology, 36,* 628-637.

O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32,* 396-402.

Pasquini, M., Picardi, A., Biondi, M., Gaetano, P., & Morosini, P. (2004). Relevance of anger and irritability in outpatients with major depressive disorder. *Psychopathology, 37,* 155-160.

Paul, G.L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31,* 109-118.

Perlis, R.H., Fava, M., Trivedi, M.H., Alpert, J., Luther, J.F., Wisniewski, S.R. et al. (2009). Irritability is associated with anxiety and greater severity, but not bipolar spectrum features, in major depressive disorder. *Acta Psychiatrica Scandinavica, 119,* 282-289.

Piaget, J. (1962). The stages of the intellectual development of the child. *Bulletin of the Menninger Clinic, 26,* 120-128.

Picardi, A., Morosini, P., Gaetano, P., Pasquini, M., & Biondi, M. (2004). Higher levels of anger and aggressiveness in major depressive disorder than in anxiety and somatoform disorders. *Journal of Clinical Psychiatry, 65,* 442-443.

Pirkola, S., Saarni, S., Suvisaari, J., Elovainio, M., Partonen, T., Aalto, A.M. et al. (2003). General health and quality-of-life measures in active, recent, and comorbid mental disorders: a population-based health 2000 study. *Comprehensive Psychiatry, 50,* 108-114.

Reidy, J. & Keogh, E. (1997). Testing the discriminant and convergent validity of the mood and anxiety symptoms questionnaire using a British sample. *Personality and Individual Differences, 23,* 337-344.

Rettew, D.C., Lynch, A.D., Achenbach, T.M., Dumenci, L., & Ivanova, M.Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18,* 169-184.

Roca, M., Gili, M., Garcia-Garcia, M., Salva, J., Vives, M., Garcia, C.J. et al. (2009). Prevalence and comorbidity of common mental disorders in primary care. *Journal of Affective Disorders, 119,* 52-58.

Russell, J.A. & Carroll, J.M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125,* 3-30.

Russell, J.A. & Lewicka, M. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology, 57,* 848-856.

Satorra, A. & Bentler, P.M. (1988). *Scaling Corrections for Chi-square Statistics in Covariance Structure Analysis. Proceedings of the Business and Economic Statistics Section.* Alexandria, VA: American Statistical Association.

Shankman, S.A. & Klein, D.N. (2003). The relation between depression and anxiety: an evaluation of the tripartite, approach-withdrawal and valence-arousal models. *Clinical Psychology Review, 23,* 605-637.

Shear, M.K., Brown, T.A., Barlow, D.H., Money, R., Sholomskas, D.E., Woods, S.W. et al. (1997). Multicenter collaborative panic disorder severity scale. *American Journal of Psychiatry, 154,* 1571-1575.

Sheehan, D.V., Janavs, J., Baker, R., Harnett-Sheehan, K., Knapp, E., Sheehan, M. et al. (1998). MINI - Mini International Neuropsychiatric Interview - English Version 5.0.0 - DSM-IV. *Journal of Clinical Psychiatry, 59,* 34-57.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E. et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59 Suppl 20,* 22-33.

Silverstone, P.H. & Von Studnitz E. (2003). Defining anxious depression: going beyond comorbidity. *Canadian Journal of Psychiatry.Revue Canadienne de Psychiatrie, 48,* 675-680.

Simms, L.J., Gros, D.F., Watson, D., & O'Hara, M.W. (2008). Parsing the general and specific components of depression and anxiety with bifactor modeling. *Depression and Anxiety, 25,* E34-46.

Simms, L.J., Prisciandaro, J.J., Krueger, R.F., & Goldberg, D.P. (2012). The structure of depression, anxiety and somatic symptoms in primary care. *Psychological Medicine, 42,* 15-28.

Slade, M. (2002). Routine outcome assessment in mental health services. *Psychological Medicine, 32,* 1339-1343.

Slade, M., McCrone, P., Kuipers, E., Leese, M., Cahill, S., Parabiaghi, A. et al. (2006). Use of standardised outcome measures in adult mental health services: randomised controlled trial. *British Journal of Psychiatry, 189,* 330-336.

Spector, P.E., Van Katwyk, P.T., Brannick, M.T., & Chen, P.Y. (1997). When Two Factors Don't Reflect Two Constructs: How Item Characteristics Can Produce Artifactual Factors. *Journal of Management, 23,* 659-677.

Sperry, L., Brill, P.L., Howard, K.I., & Grissom, G.R. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions*. Philadelphia, PA: Brunner/Mazel.

Stiles, W.B., Barkham, M., Twigg, E., Mellor-Clark, J., & Cooper, M. (2006). Effectiveness of cognitive-behavioural, person-centred and psychodynamic therapies as practised in UK National Health Service settings. *Psychological Medicine, 36,* 555-566.

Tellegen, A., Watson, D., & Clark, L.A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10,* 297-303.

Tomarken, A.J., Dichter, G.S., Freid, C., Addington, S., & Shelton, R.C. (2004). Assessing the effects of bupropion SR on mood dimensions of depression. *Journal of Affective Disorders, 78,* 235-241.

Tyrer, P. (2001). The case for cothymia: mixed anxiety and depression as a single diagnosis. *British Journal of Psychiatry, 179,* 191-193.

Tyrer, P., Owen, R.T., & Cicchetti, D.V. (1984). The brief scale for anxiety: a subdivision of the comprehensive psychopathological rating scale. *Journal of Neurology, Neurosurgery and Psychiatry, 47,* 970-975.

Van Kampen, D. (2006). The Dutch DAPP-BQ: improvements, lower- and higher-order dimensions, and relationship with the 5DPT. *Journal of Personality Disorders, 20,* 81-101.

Van Kampen, D., De Beurs, E., & Andrea, H. (2008). A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): the DAPP-SF. *Psychiatric Research, 160,* 115-128.

Van Noorden, M.S., Giltay, E.J., Den Hollander-Gijsman, M.E., van der Wee, N.J., Van Veen, T., & Zitman, F. G. (2010). Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: the Leiden Routine Outcome Monitoring Study. *Journal of Affective Disorders, 125,* 116-123.

Van Veen, T., Goeman, J.J., Monajemi, R., Wardenaar, K.J., Hartman, C.A., Snieder, H. et al. (2012). Different gene sets contribute to different symptom dimensions of depression and anxiety. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics, 159B,* 519-528.

Van Veen, T., Wardenaar, K.J., Carlier, I.V., Spinhoven, P., Penninx, B.W., & Zitman, F.G. (2013). Are childhood and adult life adversities differentially associated with specific symptom dimensions of depression and anxiety? Testing the tripartite model. *Journal of Affective Disorders, 146,* 238-245.

Van Vliet, I.M., Leroy, H., & van Megen, H.J.G.M. (2000). *M.I.N.I. International Neuropsychiatric Interview; Dutch version 5.0.0.-R*. Department of Psychiatry University Medical Center Utrecht, the Netherlands.

Van Widenfelt, B.M., Treffers, P.D., De Beurs, E., Siebelink, B.M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review, 8,* 135-147.

Veen, G., Giltay, E.J., De Rijk, R.H., Van Vliet, I.M., Van Pelt, J., & Zitman, F.G. (2009). Salivary cortisol, serum lipids, and adiposity in patients with depressive and anxiety disorders. *Metabolism, 58,* 821-827.

Veen, G., Van Vliet, I.M., De Rijk, R.H., Giltay, E.J., Van Pelt, J., & Zitman, F.G. (2011). Basal cortisol levels in relation to dimensions and DSM-IV categories of depression and anxiety. *Psychiatric Research, 185,* 121-128.

Vieta, E., Reinares, M., & Rosa, A.R. (2011). Staging bipolar disorder. *Neurotoxicity Research, 19,* 279-285.

Vollebergh, W.A.M., Iedema, J., Bijl, R.V., De Graaf, R., Smit, F., & Ormel, J. (2001). The structure and stability of common mental disorders-the NEMESIS Study. *Archives of General Psychiatry, 58,* 597-603.

Wardenaar, K.J., Giltay, E.J., Van Veen T., Zitman, F.G., & Penninx, B.W. (2012). Symptom dimensions as predictors of the two-year course of depressive and anxiety disorders. *Journal of Affective Disorders, 136,* 1198-1203.

Wardenaar, K.J., Van Veen T., Giltay, E.J., De Beurs, E., Penninx, B.W., & Zitman, F.G. (2010). Development and validation of a 30-item short adaptation of the Mood and Anxiety Symptom Questionnaire (MASQ). *Psychiatric Research, 179,* 101-106.

Wardenaar, K.J., Vreeburg, S.A., Van Veen, T., Giltay, E.J., Veen, G., Penninx, B.W. et al. (2011). Dimensions of depression and anxiety and the hypothalamo-pituitary-adrenal axis. *Biological Psychiatry, 69,* 366-373.

Ware, J. E., Jr. & Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30,* 473-483.

Watson, D. & Clark, L.A. (1991). *The Mood and Anxiety Symptom Questionnaire*. Iowa City: University of Iowa, Department of Psychology.

Watson, D., Clark, L.A., Weber, K., Assenheimer, J.S., Strauss, M.E., & McCormick, R.A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology, 104,* 15-25.

Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219-235.

Watson, D., Weber, K., Assenheimer, J.S., Clark, L.A., Strauss, M.E., & McCormick, R.A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology, 104,* 3-14.

Watson, D. (2005). Rethinking the Mood and Anxiety Disorders: A Quantitative Hierarchical Model for DSM-V. *Journal of Abnormal Psychology, 114,* 522-536.

Widiger, T.A. & Samuel, D.B. (2005). Diagnostic Categories or Dimensions? A Question for the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition. *Journal of Abnormal Psychology, 114,* 494-504.

Wing, J.K., Beevor, A.S., Curtis, R.H., Park, S.B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry, 172,* 11-18.

Wittchen, H.U., Carter, R.M., Pfister, H., Montgomery, S.A., & Kessler, R.C. (2000). Disabilities and quality of life in pure and comorbid generalized anxiety disorder and major depression in a national survey. *International Clinical Psychopharmacology, 15,* 319-328.

Wittchen, H.U., Kessler, R.C., Beesdo, K., Krause, P., Hofler, M., & Hoyer, J. (2002). Generalized anxiety and depression in primary care: prevalence, recognition, and management. *Journal of Clinical Psychiatry, 63 Suppl 8,* 24-34.

Wood, A.M., White, I.R., & Thompson, S.G. (2004). Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials, 1,* 368-376.

Wright, B.D. & Masters, G.N. (1982). *Rating Scale Analysis, Rasch Measurement*. Chicago, Il: MESA press.

Wright, B.D. & Stone, M.H. (1979). *Best Test Design: Rasch Measurement.* Chicago, Il: MESA press.

Zinbarg, R.E. & Barlow, D.H. (1996). Structure of Anxiety and the Anxiety Disorders: A Hierarchical Model. [Article]. *Journal of Abnormal Psychology, 105,* 181-193.

Zitman, F.G. (2012). Stagering, profilering en routine outcome monitoring. *Tijdschrift voor Psychiatrie, 54,* 979-984.

SAMENVATTING

Er is de laatste decennia veel onderzoek verricht naar de oorzaken en optimale behandeling van depressies en angststoornissen. De geboekte vooruitgang is echter relatief beperkt, omdat onderzoeksresultaten veelal niet consistent en soms zelf tegenstrijdig blijken.

Een veel genoemde mogelijke oorzaak hiervoor is het gebruik van de DSM-IV om de stoornissen vast te stellen. De DSM (Diagnostic and Statistical Manual of Mental Disorders) is een Amerikaans handboek met diagnoses van psychische aandoeningen dat in de meeste landen als standaard in de psychiatrische diagnostiek dient en door vrijwel alle psychiaters en psychologen gebruikt wordt. De eerste editie is verschenen in 1952 en vanaf de derde editie (DSM III, 1980) worden stoornissen duidelijk beschreven en wordt er gedefinieerd welke symptomen kunnen voorkomen bij een ziektebeeld, en hoeveel symptomen aanwezig dienen te zijn, voordat er gesproken kan worden van een bepaald syndroom of ziektebeeld bij een patiënt. Voor die tijd werden termen als 'depressie' en 'psychose' door verschillende auteurs heel anders ingevuld. Door heldere standaard beschrijvingen werd de classificatie meer betrouwbaar en kon men er ook internationaal over communiceren. Als voorbeeld staan in tekstvak 1 de criteria die in de DSM-IV gehanteerd worden voor het vaststellen van een depressieve stoornis. In de onderzoeken in dit proefschrift is gebruik gemaakt van de DSM-IV-TR (2000). Zeer recent in mei 2013 is de DSM-5 verschenen.

Naast de grote winst die de standaardisatie in diagnoses heeft gebracht, heeft deze categoriale benadering ook nadelen. Ten eerste is er veelvuldig sprake van comorbiditeit; DSM-diagnoses komen veel samen voor. Ten tweede zijn de grenzen tussen ziek en gezond (wel of geen DSM-diagnose) kunstmatig. Er is een afkappunt bepaald, maar in werkelijkheid is er geen afkappunt tussen ziek en gezond, net zoals er geen duidelijk moment aan te geven is wanneer de dag overgaat in de nacht: er is sprake van een continue verdeling.

Een derde probleem van de DSM is dat er binnen één diagnosegroep veel verschil kan zijn in symptomen. Om deze heterogeniteit te illustreren, staan in tekstvak 2 de klachten van twee fictieve patiënten beschreven, die allebei voldoen aan de DSM-criteria voor een depressieve stoornis. Hoewel beide personen dezelfde diagnose hebben, is er een groot verschil in de symptomen die zij ervaren. Als zij elkaar zouden spreken, zouden zij vast en zeker verbaasd zijn dat ze voor dezelfde ziekte behandeld (gaan) worden. Dit is mogelijk omdat het hebben van vijf van negen klachten voldoende is om aan de DSM diagnose depressie te voldoen (zie tekstvak 1). Bovendien bevatten sommige criteria zowel een toename als een afname van een bepaald gedrag (bijvoorbeeld gewichtstoename en – afname). Dit illustreert dat een groep patiënten met een depressie een heterogene groep is.

DSM-IV criteria Depressie

Vijf (of meer) van de volgende symptomen zijn binnen dezelfde periode van twee weken bijna dagelijks aanwezig en wijzen op een verandering ten opzichte van het eerdere functioneren; ten minste een van de symptomen is ofwel depressieve stemming ofwel verlies van interesse of plezier.

1. Depressieve stemming gedurende het grootste deel van de dag
2. Duidelijke vermindering van interesse of plezier in alle of bijna alle activiteiten gedurende het grootste deel van de dag
3. Duidelijke gewichtsvermindering zonder diet, of gewichtstoename of afgenomen of toegenomen eetlust
4. Insomnia of hypersomnia
5. Psychomotorische agitatie of remming
6. Moeheid of verlies van energie
7. Gevoelens van waardeloosheid of buitensporige of onterechte schuldgevoelens
8. Verminderd vermogen tot nadenken of concentratie of besluiteloosheid
9. Terugkerende gedachten aan de dood, terugkerende suïcidegedachten, of een (specifiek plan voor een) suïcidepoging

Tekstvak 1

Twee patiënten met een depressie

Patient A (vrouw, 36 jaar) voelt zich verdrietig en waardeloos, heeft moeite om in slaap te komen, en heeft gewicht verloren. Ze voelt zich erg schuldig tegenover haar gezin (man en twee jonge kinderen) en wenst vaak dat ze de volgende ochtend niet meer wakker wordt ("dat zou het beste zijn voor iedereen").

Patient B (vrouw, 29 jaar) geniet niet meer zo van haar leven als ze voorheen deed. Ze slaapt erg veel en heeft weinig energie, ze is in de afgelopen maanden 5 kilo aangekomen. Ze krijgt weinig voor elkaar en kan zich niet concentreren.

Tekstvak 2

Een dimensionele benadering lijkt de oplossing voor alle drie de problemen. Je stelt daarbij geen diagnoses meer, maar karakteriseert iemand aan de hand van de mate waarin verschillende kenmerken voorkomen. Dat doe je

bij iedereen op dezelfde manier, los van eventuele DSM diagnoses. Zo wordt bijvoorbeeld bij iedereen bepaald hoe somber hij is, hoe waardeloos hij zich voelt, hoe angstig enzovoort. De mate waarin die kenmerken voorkomen wordt bepaald aan de hand van een reeks vragen die elk gescoord moeten worden op de mate waarin ze van toepassing zijn, bijvoorbeeld op een schaal van 0 tot 5. Comorbiditeit is dan geen probleem meer, want iedereen wordt op dezelfde kenmerken 'gescoord', ook het onderscheid tussen ziek en gezond speelt niet meer en tenslotte kan met een dimensionele benadering ook beter rekening worden gehouden met de heterogeniteit.

Het doel van dit proefschrift was om de haalbaarheid van dimensionele modellen te testen in een grote steekproef psychiatrische poliklinische patiënten met stemmingsstoornissen, angst en/of somatoforme stoornissen en een dimensioneel model te ontwikkelen dat de nadelen van bestaande modellen overwint.

De databestanden met gegevens van patiënten die in dit proefschrift zijn gebruikt, zijn verzameld met Routine Outcome Monitoring (ROM). ROM is het herhaald meten van de klachten en problemen van de patiënten voorafgaand aan, tijdens en na de behandeling. In hoofdstuk 2 wordt de methode ROM in detail beschreven. De studie laat zien dat implementatie van ROM haalbaar is en dat veel behandelaren ROM beschouwen als een belangrijke aanvulling voor de klinische behandeling. Ook stimuleert ROM onderzoek; de verzamelde gegevens kunnen immers gebruikt worden voor onderzoek naar de fenomenologie van psychiatrische stoornissen (zoals dit proefschrift) en de uitkomsten van behandeling zoals gegeven in de dagelijkse klinische praktijk. Opvallend is de hoge comorbiditeit in ons patiëntenbestand. Volgens het diagnostische interview dat bij alle patiënten is afgenomen als onderdeel van ROM (de MINI-Plus), voldoen 1,618 patiënten (42.6%) aan de criteria van een enkele stemmings-, angst- of somatoforme stoornis. Een bijna even groot aantal patiënten, 1,556 patiënten (41.0%) had ten tijde van de meting meer dan één van deze stoornissen: 967 patiënten (25.5%) met twee stoornissen, 403 patiënten (10.6%) met drie, en 186 patiënten (4.9%) met vier of meer stoornissen tegelijk. Dit veelvuldig tegelijk voorkomen van verschillende van deze stoornissen laat zien dat de DSM categorieën niet uitsluitend en onderscheidend zijn en benadrukt de behoefte aan een dimensioneel model.

In hoofdstuk 3 hebben we de relatie onderzocht tussen ernst en comorbiditeit van depressie en angststoornissen in een groot patiënten bestand. Zoals verwacht, was de ernst van symptomen hoger in de groep met comorbiditeit vergeleken met de groepen met slechts één stoornis (geen comorbiditeit). De patiënten met meer dan één stoornis waren er dus slechter aan toe dan de patiënten met een enkele stoornis. De ernst van de depressieve symptomen in de comorbide groep was hoger dan in de groep met een 'pure' depressie en de ernst van angstsymptomen in de comorbide groep was hoger dan in de

groep met een 'pure' angststoornis. Opmerkelijk was ook dat de gemiddelde score op de angstschalen niet significant verschilde tussen patiënten met een pure depressie en patiënten met een pure angststoornis. Bevindingen zoals hierboven beschreven blijven verborgen als alleen categoriaal wordt gediagnosticeerd. Bij een dimensionele benadering komen ze wel aan het licht.

Een veel gebruikt dimensioneel model om de symptomen van depressie en angststoornissen te beschrijven is het 'tripartite model', ontwikkeld door Clark en Watson in 1991. Dit model bestaat uit drie dimensies en is ontwikkeld om de gemeenschappelijke en specifieke symptoom domeinen van depressie en angst te berschrijven. De eerste dimensie *Negatief Affect* (NA) bestaat uit symptomen van algemeen psychisch onwelbevinden, zoals boos, ontdaan, en verdrietig. NA is een algemene factor en is niet specifiek voor depressie of angst. De twee andere factoren in het model zijn respectievelijk specifiek voor depressie en angst. De dimensie *Positief Affect* (PA) bestaat uit symptomen van gebrek aan interesse en energie, en is specifiek voor depressie. De dimensie *Somatic Arousal* (SA) bestaat uit symptomen van somatische (lichamelijke) opwinding, zoals versnelde hartslag en zweten. Het is specifiek voor angst.

De drie dimensies van het tripartite model kunnen gemeten worden met een vragenlijst; de MASQ (Mood and Anxiety Symptom Questionnaire). Om onderzoek te kunnen doen naar het tripartite model in Nederlandse databestanden, was een vertaling van de MASQ een vereiste. In hoofdstuk 4 wordt de Nederlandse vertaling van de MASQ gepresenteerd en de toepasbaarheid en validiteit getest. De psychometrische kwaliteiten van de Nederlandse MASQ bleken ruim voldoende te zijn. In overeenstemming met het model vonden we drie schalen terug in de data, die bovendien goed onderscheid kunnen maken tussen subgroepen van patiënten met stemmings- en angststoornissen.

Hoewel het tripartite model veel onderzoek naar dimensies heeft gestimuleerd, heeft het ook kritiek ontvangen. Een belangrijk punt van kritiek is dat de symptomen van de verschillende angststoornissen geen deel uitmaken van het model. De dimensie SA beslaat niet alle angststoornissen, maar vooral de paniek stoornis. Het is een nauwe conceptualisatie van angst. Het beperkt zich tot de lichamelijke uitingen van angst en negeert andere belangrijke aspecten zoals piekeren en vermijdings- oftewel fobische angst.

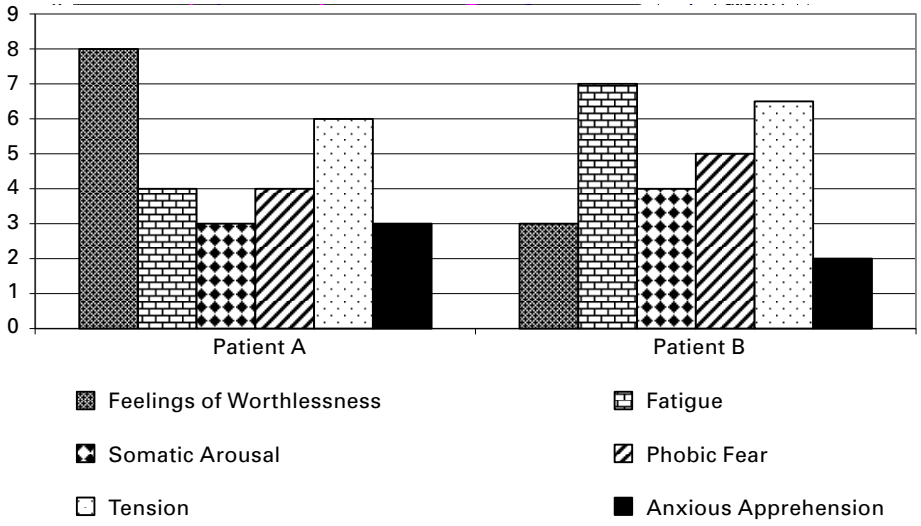
In hoofdstuk 5 presenteren we een eerste eigen dimensioneel model dat een aantal belangrijke aspecten van depressie en angst omvat en bestaat uit duidelijk van elkaar verschillende dimensies. Ons doel was onder andere om meer symptomen van angst in het nieuwe model op te nemen dan was gebeurd in het tripartite model. We hebben in dit onderzoek vragen gebruikt van 2 vragenlijsten: de Mood and Anxiety Symptom Questionnaire (MASQ, Watson & Clark, 1991) en de Brief Symptom Inventory (BSI, Derogatis, 1975). Het gepresenteerde model bestaat uit vijf dimensies: sombere stemming, laag

positief affect, somatische arousal, fobische angst en vijandigheid (depressed mood, lack of positive affect, somatic arousal, phobic fear, hostility).

De validiteit van het model werd ondersteund door de volgende bevindingen: de schalen bleken in staat om onderscheid te maken tussen patiënten met ofwel een stemmings- ofwel een angststoornis. Laag positief affect en fobische angst onderscheidden het best tussen depressieve patiënten en patiënten met een angststoornis. Ook lieten de verschillende angststoornissen een verschillend patroon zien qua scoring op de nieuwe dimensies. Binnen de angststoornissen was somatische arousal specifiek voor patiënten met een paniekstoornis. Fobische angst bleek geassocieerd met een paniekstoornis, een specifieke fobie en een sociale angststoornis, maar niet met een gegeneraliseerde angststoornis (GAD).

Overwegende dat de in hoofdstuk 5 beschreven studie het model valideert door vergelijking van de dimensies met de categorieën van de DSM-IV (stoornis-gebaseerde aanpak), hebben we in hoofdstuk 6 een stap gedaan verder weg van de DSM-IV en meer in de richting van een 'echt dimensioneel model '(symptoom-gebaseerde aanpak). Als uitgangspunt hebben we gebruik gemaakt van een groot aantal items (vragen in vragenlijst): (1) De items van de Mood and Anxiety Symptom Questionnaire (Watson & Clark, 1991) om NA, PA en SA te meten, (2) Items van de angst subschalen van de Brief Symptom Inventory (Derogatis, 1975), om angst te meten en (3) door onszelf geformuleerde vragen om 'anxious apprehension', piekeren over wat komen gaat, te meten. We legden de vragen voor aan 1333 poliklinische patiënten en analyseerden ze vervolgens met geavanceerde technieken om de dimensies vast te stellen. We gebruikten eerst de ingevulde vragenlijsten van een deel van de patiënten om vast te stellen om welke factoren of dimensies het gaat. Daarna keken we in de gegevens van de overige patiënten of het geen toevallige bevinding was, of we in die tweede groep hetzelfde vonden. We kwamen uit op een 6-factor model met de volgende dimensies: gevoelens van waardeloosheid, vermoeidheid, somatische arousal, angstige bezorgdheid, fobische angst en spanning (feelings of worthlessness, fatigue, somatic arousal, anxious apprehension, phobic fear, tension) . Somatische arousal, angstige bezorgdheid en fobische angst zijn allemaal duidelijk angstachtige klachten. In plaats van alleen de enkele angstdimensie van het tripartite model (somatische arousal), onderscheidt het in hoofdstuk 6 beschreven model drie groepen van angstsymptomen. Met dit dimensionele model kunnen verschillende symptomen van depressie en angst geregistreerd worden. Het doet recht aan de heterogeniteit van angst en bestaat uit eenvoudig te meten dimensies.

Het model biedt een oplossing voor de drie problemen die aan de categoriale benadering kleven. Om te beginnen maakt het de heterogeniteit in klachten en symptomen zichtbaar bij mensen die volgens de DSM dezelfde diagnose hebben. Figuur 1 geeft een grafische voorstelling van het symptoomprofiel van



Figuur 1 Voorbeeld van een 'symptoomprofiel' voor twee patiënten, gediagnosticeerd met een depressie volgens de DSM-IV.

de twee hierboven besproken fictieve patiënten die volgens de DSM-IV beiden aan een depressie lijden (zie tekstvak 2). Het is duidelijk zichtbaar dat ondanks het feit dat beide patiënten de diagnose 'depressieve stoornis' hebben, ze, als we ze karakteriseren met behulp van de zes dimensies die we gevonden hebben, een heel verschillend symptoomprofiel hebben.

Ook patiënten die volgens de DSM-IV een angststoornis hebben, kunnen worden gekarakteriseerd met dezelfde zes dimensies en zullen over het algemeen een ander profiel hebben. De aanwezige angstklachten die wij vonden bij gebruik van de DSM-IV, bij puur depressieve patiënten (zie hoofdstuk 3) zal in deze dimensionele profielen ook zichtbaar worden. Ook de klachten van patiënten met comorbiditeit kunnen met deze zes dimensies worden gekarakteriseerd. Tenslotte speelt ook de kunstmatige scheiding tussen ziek en gezond bij deze dimensies geen rol meer: iedereen die de vragenlijst invult krijgt een profiel op basis van dezelfde zes dimensies. Dit geldt ook voor mensen die geen klachten hebben en voor mensen die wel psychische klachten hebben maar volgens de DSM-IV niet aan de criteria voor een depressie of een angststoornis voldoen.

Alvorens een dimensionele benadering op grote schaal te kunnen gaan gebruiken, moet de superioriteit van de dimensionele benadering ten opzichte van de DSM-IV worden aangetoond. Niet alleen voor wat betreft het onderzoek

naar de etiologie (ontstaansgeschiedenis) van psychische stoornissen maar ook naar de klinische bruikbaarheid. Toekomstig onderzoek moet eerst uitwijzen of een dimensioneel profiel inderdaad efficiënter is om te beslissen wat het belangrijkste doelwit voor de behandeling zou moeten zijn en wat voor soort behandeling wordt aanbevolen voor een patiënt. Op termijn kan het hopelijk ook bijdragen aan het ontwikkelen van nieuwe, meer op de specifieke kenmerken van de klachten van de individuele patiënt, afgestemde behandelingen en daarmee op wat men wel noemt 'personalized medicine'. Wij zijn van mening dat de nadruk in onderzoek de komende jaren moet liggen op het bepalen van de factoren die kunnen voorspellen hoe symptomen zich ontwikkelen in de tijd en factoren die effectieve behandeling kunnen voorspellen. Een dimensionele benadering moet hier een onderdeel van zijn, en de verwachting is dat er grotere voortgang gemaakt zal worden in de kennis over depressie en angst, dan met de traditionele categoriale benadering van psychische klachten.

CURRICULUM VITAE

Margien Elisabeth den Hollander-Gijsman was born on August 19th, 1970 in Leiden, the Netherlands and she has lived in Leiden ever since. She is married to Pancras S.G. den Hollander and mother of Pancras J. den Hollander (1998) and Dorien E. den Hollander (2000).

After receiving her VWO diploma at the 'Stedelijk Gymnasium Leiden' in 1988, she studied Psychology at the University of Leiden. She obtained her Masters Degree in Social Psychology in 1995. In February 1996 she started working at TNO Prevention and Health and contributed to several research projects and publications.

From 2002 until 2010 she fulfilled several functions at the Department of Psychiatry of the Leiden University Medical Center (LUMC). The first two years she was involved with the implementation of Routine Outcome Monitoring (ROM). In September 2004 she started with her PhD-project and from January 2007 to March 2010 she combined her research work with the further optimalization of ROM both at the LUMC and Rivierduinen.

From March 2010 until August 2013 she was employed at NETQ Healthcare; provider of NETQ ROM – software for Routine Outcome Monitoring. At the time of the appearance of this thesis, she is ready for a new challenge.

LIST OF PUBLICATIONS

De Beurs, E., Den Hollander-Gijsman, M.E., Buwalda, V., Trijsburg, W., & Zitman, F.G. (2005). De Outcome Questionnaire (OQ-45): een meetinstrument voor meer dan alleen psychische klachten. *De Psycholoog, 40,* 393-400.

De Beurs, E., Den Hollander-Gijsman, M.E., Helmich, S., & Zitman, F.G. (2007). 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, 45,* 1609-1617.

De Beurs, E., Den Hollander-Gijsman, M.E., Van Rood, Y.R., Van der Wee, N.J.A., Giltay, E.J., Van Noorden, M.S. et al. (2011). Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology and Psychotherapy, 18,* 1-12.

De Beurs, E., Den Hollander-Gijsman, M.E., Van Rood, Y.R., & Zitman, F.G. (2011). The Dutch version of the panic disorder severity scale: reliability and validity. In J.K.Levin & A. D. Torterolo (Eds.), *Panic Disorder: Symptoms, Treatment and Prevention* (New York: Nova Science Publishers.

De Klerk, S., Van Noorden, M.S., van Giezen, A.E., Spinhoven, P., Den Hollander-Gijsman, M.E., Giltay, E.J. et al. (2011). Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: the Leiden Routine Outcome Monitoring study. *Journal of Affective Disorders, 133,* 257-264.

De Roos, C., Greenwald, R., Den Hollander-Gijsman, M.E., Noorthoorn, E., Van Buuren, S., & De Jongh, A. (2011). A randomised comparison of cognitive behavioural therapy (CBT) and eye movement desensitisation and reprocessing (EMDR) in disaster-exposed children. *European Journal of Psychotraumatology, 2,* 5694 - DOI: 10.3402/ejpt.v2i0.5694.

De Roos, C., Veenstra, A.C., De Jongh, A., Den Hollander-Gijsman, M.E., Van der Wee, N.J.A., Zitman, F.G. et al. (2010). Treatment of chronic phantom limb pain using a trauma-focused psychological approach. *Pain Research & Management, 15,* 65-71.

Den Hollander-Gijsman, M.E., De Beurs, E., Van der Wee, N.J.A., Van Rood, Y.R., & Zitman, F.G. (2010). Distinguishing between depression and anxiety: a proposal for an extension of the tripartite model. *European Psychiatry, 25,* 197-205.

Den Hollander-Gijsman, M.E., Wardenaar, K.J., De Beurs, E., Van der Wee, N.J., Mooijaart, A., Van Buuren, S. et al. (2012). Distinguishing symptom dimensions of depression and anxiety: an integrative approach. *Journal of Affective Disorders, 136,* 693-701.

Smits, N., Zitman, F.G., Cuijpers, P., Den Hollander-Gijsman, M.E., & Carlier, I.V. (2012). A proof of principle for using adaptive testing in routine Outcome Monitoring: the efficiency of the Mood and Anxiety Symptoms Questionnaire -Anhedonic Depression CAT. *BMC Medical Research Methodology, 12,* doi: 10.1186/1471-2288-12-4.

Van der Meer, J., Van Rood, Y.R., Van der Wee, N.J.A., Den Hollander-Gijsman, M.E., Van Noorden, M.S., Giltay, E.J. et al. (2012). Prevalence, demographic and clinical characteristics of body dysmorphic disorder among psychiatric outpatients with mood, anxiety or somatoform disorders. *Nordic Journal of Psychiatry, 66,* 232-238.

Van Fenema, E.M., Van der Wee, N.J., Giltay, E.J., Den Hollander-Gijsman, M.E., & Zitman, F. G. (2012). Vitality predicts level of guideline-concordant care in routine treatment of mood, anxiety and somatoform disorders. *Journal of Evaluation in Clinical Practice, 18,* 441-448.

Van Noorden, M.S., Giltay, E.J., Den Hollander-Gijsman, M.E., van der Wee, N.J., Van Veen, T., & Zitman, F. G. (2010). Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: the Leiden Routine Outcome Monitoring Study. *Journal of Affective Disorders, 125,* 116-123.

Van Noorden, M.S., Minkenberg, S.E., Giltay, E.J., Den Hollander-Gijsman, M.E., Van Rood, Y.R., van der Wee, N.J. et al. (2011). Pre-adult versus adult onset major depressive disorder in a naturalistic patient sample: the Leiden Routine Outcome Monitoring Study. *Psychological Medicine, 41,* 1407-1417.

Vogels, T., Reijneveld, S.A., Brugman, E., Den Hollander-Gijsman, M.E., Verhulst, F.C., & Verloove-Vanhorick, S. P. (2003). Detecting psychosocial problems among 5-6-year-old children in preventive Child Health Care: the validity of a short questionnaire used in an assessment procedure for detecting psychosocial problems among children. *European Journal of Public Health, 13,* 353-360.

Wardenaar, K.J., Van Veen, T., Giltay, E.J., Den Hollander-Gijsman, M.E., Penninx, B.W., & Zitman, F.G. (2010). The structure and dimensionality of the Inventory of Depressive Symptomatology Self Report (IDS-SR) in patients with depressive disorders and healthy controls. *Journal of Affective Disorders, 125,* 146-154.

DANKWOORD

Bij de totstandkoming van dit proefschrift zijn veel mensen betrokken. Ik bedank mijn promotor, Frans Zitman, voor alle mogelijkheden die hij mij geboden heeft om mijzelf te ontwikkelen in verschillende functies binnen het LUMC. Het is een voorrecht om ook na zijn emeritaat met hem samen aan dit proefschrift te hebben mogen werken. Mijn copromotor, Edwin de Beurs, dank ik voor zijn begeleiding. Van hem heb ik geleerd dat je niet altijd van tevoren op alle mogelijke tegenslagen voorbereid hoeft te zijn; soms is het beter gewoon te beginnen. Nic van der Wee bedank ik voor zijn bereidheid om begin 2009 toe te treden tot de promotiecommissie, en voor zijn waardevolle bijdragen die daarop volgden.

Alle data waarop dit proefschrift is gebaseerd zijn verzameld dankzij de implementatie van Routine Outcome Monitoring (ROM) in het LUMC en Rivierduinen. Naast het directe nut voor de klinische praktijk, maakt ROM grootschalig onderzoek mogelijk in de dagelijkse praktijk van de GGZ. Ik bedank de Raden van Bestuur van het LUMC en Rivierduinen die dit project mogelijk hebben gemaakt. Veel dank is verschuldigd aan alle testmedewerkers ROM van de afdeling Psychiatrie van het LUMC en van Rivierduinen. Zij motiveren zowel patiënt als behandelaar door het nut van ROM zichtbaar te maken, en zorgen voor betrouwbare gegevens. Zonder hen waren de waardevolle databestanden niet tot stand gekomen.

In 2002 ben ik in een duobaan met Marieke de Wit gestart met Routine Outcome Monitoring in het LUMC. Ik dank haar voor de fijne samenwerking, haar eerlijkheid en interesse.

Data verzamelen is één ding, data-schoning en -analyse is een tweede. Met veel plezier denk ik terug aan de ROM-baseline club. In het bijzonder wil ik Yanda van Rood, Martijn van Noorden en Erik Giltay bedanken voor hun samenwerking, hulpvaardigheid en luisterend oor. Ook de samenwerking met Klaas Wardenaar heb ik als zeer prettig ervaren. Ik wil hem hartelijk danken voor zijn werk aan hoofdstuk 6, zijn nuchtere blik en attente berichten als ik iets te lang niets had laten horen.

Jarenlang heb ik gewerkt aan een bureau in de kantoortuin van de afdeling Psychiatrie in het LUMC. Vele onderzoekers heb ik zien komen en gaan. Ik dank hen allen voor hun bijdrage aan het goede wetenschappelijke klimaat. Blij werd ik van de enthousiaste ontvangst wanneer ik in de laatste schrijffase een dagje in de kantoortuin kwam werken als mijn agenda dat toeliet.

Ik bedank medeonderzoekers Frederieke van Veen, Ellen Klaassens, Marie-José van Tol en Floriana Luppino voor hun vriendschap. Mijn vriendinnen Corine van Lavieren, Tessy Geurink en Nathalie Vriezelaar bedank ik voor de afleiding die zij mij brachten, vaak tot in de kleine uurtjes. Ook mijn vriendin Hendrien Kaal heeft mij altijd gesteund, ondanks haar waarschuwende woorden toen ik

met het promotie-traject ging starten.

Mijn paranimfen, Ellen en Tessy, ben ik zeer erkentelijk voor het feit dat zij aan mijn zijde staan in het Academiegebouw en ook daarbuiten. Mijn ouders ben ik zeer dankbaar voor alle kansen en aanmoedigingen die zij mij geven. Pancras, wil ik met heel mijn hart bedanken voor het geduld en de ruimte die hij mij gegeven heeft om mijn promotietraject te voltooien. Tegen onze kinderen Pancras jr. en Dorien zeg ik: “Hoera, mamma’s werkstuk is klaar!”

Stellingen behorend bij het proefschrift

OUT OF THE BOX

Moving from categories to dimensions
in de phenomenology of depression and anxiety

1. Het gebruik van symptoomprofielen doet meer recht aan verschillen tussen individuele patiënten dan het gebruik van een categoriale indeling (dit proefschrift).
2. De bevinding dat patiënten met een depressie (zonder comorbide angststoornis) even hoog scoren op angstschalen, als patiënten met een angststoornis, bewijst hoe complex de relatie tussen depressie en angststoornissen is (dit proefschrift).
3. Een gegeneraliseerde angststoornis lijkt meer op een depressie dan op de andere angststoornissen in de DSM-IV (dit proefschrift).
4. Ondanks de hoge comorbiditeit tussen depressie en angststoornissen, is het mogelijk unidimensionele constructen aan te wijzen die specifiek zijn voor de verschillende stoornissen (dit proefschrift).
5. Longitudinale data die in de dagelijkse praktijk routinematig worden verzameld zijn meer geschikt voor hypothese vormend onderzoek dan voor hypothese toetsend onderzoek.
6. Voor het succesvol implementeren van Routine Outcome Monitoring is inzicht in het zorgproces binnen de organisatie de eerste voorwaarde.
7. Wij begrijpen mensen en gebeurtenissen vaak door hen in categorieën te plaatsen. Op die manier vereenvoudigt het categoriseren ons dagelijks leven. Maar categoriseren leidt ook vaak tot fouten doordat er details weggelaten worden, of doordat er misinformatie wordt toegevoegd (L. Berkowitz, sociaal psycholoog, 1926).
8. Classificatie is de dood van de fenomenologie (K. Jaspers, Duits filosoof 1883-1969).
9. Ten gevolge van zijn persoonlijkheid en zijn omstandigheden, leeft ieder, zonder uitzondering, in een zekere beperktheid van begrippen en opvattingen. (A. Schopenhauer, Duits filosoof 1788-1860).
10. De productiviteit op een kantoortuin met onderzoekers is gerelateerd aan het aantal bezette bureaus; Het verband heeft de vorm van een omkeerde U.
11. "Strakjes" duurt soms heel lang (Pancras jr. en Dorien den Hollander, 2004).