

Endorsement frequencies and factor structure of DSM-III-R and DSM-IV Generalized Anxiety Disorder symptoms in women: implications for future research, classification, clinical practice and comorbidity

THOMAS S. KUBARYCH,¹ STEVEN H. AGGEN,¹ JOHN M. HETTEMA,¹ KENNETH S. KENDLER,^{1, 2} MICHAEL C. NEALE^{1, 2}

¹ Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia, USA

² Department of Human Genetics, Virginia Commonwealth University, Richmond, Virginia, USA

Abstract

We investigated dimensions of liability to Generalized Anxiety Disorder (GAD) and whether evidence exists for distinct pathological versus normal clusters in the population. Structured interviews were administered to a general population sample of 2,163 female twins in a cross-sectional design. Endorsement rates were estimated using full information maximum likelihood factor analyses of the DSM-III-R and DSM-IV GAD symptoms, which provides appropriate treatment of the stem-probe structure of the clinical interview.

Endorsement rates were highest for symptoms retained in DSM-IV. For both DSM-III-R and DSM-IV, a two-factor model fit the data better than a single-factor model. There was no evidence for non-normality in the liability to GAD. For DSM-III-R, autonomic symptoms loaded on a factor with panic disorder, while fatigability, difficulty concentrating and hypervigilance loaded on a factor with major depression. For DSM-IV, all items loaded on one factor, and muscle tension also loaded on a second. Major depression, panic, phobias and alcohol dependence diagnoses also loaded on the first factor. Conclusions: future research involving structured interviews should take into account the stem-and-probe format and focus on common factors rather than separate disorders; GAD is not a unidimensional construct and pathological anxiety may differ only quantitatively from normal anxiety.

Key words: Generalized Anxiety Disorder, factor analysis, missing data, structured interview, discriminant validity

Introduction

Overlap between the symptoms of Generalized Anxiety Disorder (GAD), the other anxiety disorders and depression causes considerable difficulty in differential diagnosis. Since GAD's introduction in DSM-III (American Psychiatric Association, 1980), revisions in associated symptom criteria, required duration, and diagnostic hierarchy of the disorder have largely been motivated by the need to address issues of prevalence, reliability, discriminant validity and differences between

normal and pathological worry. DSM-III-R (American Psychiatric Association, 1987) criteria for a diagnosis of GAD required at least 6 months of unrealistic or excessive worry about a variety of circumstances unrelated to another Axis I disorder. Generalized Anxiety Disorder was assessed in DSM-III-R by 18 symptoms divided into three clusters: motor tension, autonomic hyperactivity and vigilance and scanning. Six or more symptoms, experienced concurrently, were required for the diagnosis.

We are not aware of any studies that address the structure of GAD symptoms in the general population. Kenardy, Evans and Oei (1992) conducted a principal components analysis of 390 outpatients with anxiety disorders on patients' ratings on the Anxiety Symptoms and Beliefs Scale. A four-component solution emerged. The components were interpreted as respiratory symptoms, vestibular symptoms, autonomic arousal and psychological threat. This analysis was conducted on a self-selected sample of patients with a variety of anxiety disorders seeking treatment.

Merikangas, Zhang, Avenevoli, Acharyya, Neuenschwander and Angst (2003) argued that persistence of subthreshold-level anxiety and depression from early adulthood to mid-adulthood illustrates the importance of studying the continuum of anxiety and depression, not just diagnostic cases. These investigators found that, in a community-based cohort, comorbid anxiety and depression were more persistent than either syndrome alone. Anxiety alone tended to develop into either depression alone or comorbid anxiety and depression over time. Depression alone and depression comorbid with anxiety tended to be more stable than anxiety alone over time. The patterns were similar for subthreshold and threshold disorders.

Whether or not a diagnosis is warranted is a subtly different question from that of whether or not a person is suffering from symptoms. Merikangas et al. (2003) emphasize that contemporary diagnostic systems fail to cover depressive and anxiety states among subjects who do not meet duration and impairment criteria yet exhibit recurrent distress and have a history of treatment. There are fluctuations across threshold and subthreshold levels over time, so classification only by means of threshold level criteria at each evaluation will fail to capture the majority of cases with persistent anxiety through the lifespan.

The reduction to six symptom criteria in DSM-IV (American Psychiatric Association, 1994) was primarily based on a single multi-site study that assessed the frequency and utility of the 18 DSM-III-R GAD symptoms (Marten, Brown, Barlow, Borkovec, Shear, and Lydiard, 1993). This study of 204 GAD patients at four sites identified seven symptoms as having satisfactory reliability and endorsement frequency: irritability, restlessness, muscle tension, difficulty concentrating, sleep difficulties, feeling keyed up and easy fatigability. Two of these symptoms, restlessness and feeling keyed up, were combined into one criterion in DSM-IV. None of

these symptoms belonged to the autonomic hyperactivity cluster.

One of the goals in reducing the number of criteria symptoms for GAD diagnosis was to improve discriminant validity. Turvey, Stevens and Merikangas (1999) compared the discriminant validity of DSM-III-R versus DSM-IV criteria for GAD using data from a study of the familial aggregation of anxiety disorders and alcoholism. No differences were found in the discriminant validity of the two definitions. The authors speculated that the lack of difference in validity between the two definitions was due to low prevalence of autonomic hyperactivity symptoms in both their sample and earlier studies.

In a non-clinical sample of 183 students, Joorman and Stöber (1999) found that, among the six DSM-IV GAD symptoms, only muscle tension was uniquely related to pathological worry. Difficulty concentrating was related to depression. The authors concluded that muscle tension is a specific aspect of pathological worry and that discriminant validity between GAD and major depression could be improved by criteria that emphasize muscle tension and de-emphasize difficulty concentrating. Ruscio, Borkovec and Ruscio (2001), however, suggested that pathological worry differs quantitatively rather than qualitatively from normal worry.

The DSM is designed to classify subjects as either cases or non-cases. The question of whether psychopathology is better viewed as discrete classes or as a continuous dimension of liability is much discussed (Waller and Meehl, 1998; Loranger, 1999; Clark, 1999; Costa and Widiger, 2002; Pickles and Angold, 2003). There is a substantial loss of information, and consequently statistical power, when a continuous variable is treated as discrete. For example, even under optimal conditions of 50% prevalence, approximately three times the sample size is needed to assess familial resemblance for a dichotomized rather than continuous variable (Neale, Eaves and Kendler, 1994).

In the present study, we investigate the item properties and factor structure of the 18 DSM-III-R and 6 DSM-IV GAD symptomatic criteria in a population sample. Our aims are:

- to conduct factor analyses of the SCID-based DSM-III-R and IV stem items and probe symptoms in a general population sample taking into consideration the stem-and-probe format;

- to investigate the dimensionality of the DSM-IV symptoms;
- to test for non-normality in these dimensions, which would indicate that subjects' liabilities on that factor were not distributed along a continuum and suggest different categories; and
- to identify factors in GAD symptoms that may be differentially associated with various patterns of comorbidity, treatment efficacy and outcomes.

Method

Subjects were 2,163 Caucasian same-sex female twins from the first wave of the population-based Virginia Twin Registry. The response rate was 92%. The age range was 17 to 54 years (mean 30.1, SD 7.6), and median income was in the \$30,000 to \$34,999 range. Details of these samples are given by Kendler and Prescott (1999). The assessment of GAD symptoms over the previous year was based on the Structured Clinical Interview for DSM (SCID) interview for DSM-III-R (Spitzer and Williams, 1985). Two 'stem' questions were used:

- 'In the last year, have you had a time lasting at least 5 days, when you felt anxious or worried most of the time?'
- 'In the last year, have you had a time lasting at least 5 days, when most of the time your muscles felt tense, or you felt jumpy or shaky inside?'

If the response to either of these items was 'yes', the 18 DSM-III-R individual GAD symptoms ('probes') were assessed. Note that the duration requirement for these stem questions was deliberately set much lower than that required for GAD diagnosis in either DSM-III-R or DSM-IV. Empirical data support the hypothesis that shorter durations of GAD symptoms reflect the same continuum of liability as the fully syndromal disorder (Kendler, Neale, Kessler, Heath and Eaves, 1992). Nine hundred of 2,163 subjects received the 18 probes.

Factor analysis of item-level data obtained from an instrument that has a stem-probe format presents non-trivial complications. Most cases (1,263 of 2,163) in the present sample contain missing data, and would be eliminated by methods that use listwise deletion. This elimination would restrict the analysis to those with at least one positive screen and could bias the results.

The probe symptom data missing due to values of the stem items are assumed to be 'missing at random'

(MAR) according to the Little and Rubin (1987) definition. Therefore, we used full information maximum likelihood analysis of the item data to recover asymptotically unbiased estimates of the population parameters (Enders, 2001). The model used incorporates the assumption that for each item there is a normal distribution of liability underlying the response, with a threshold beyond which the item is responded to in the affirmative. This liability-threshold model is widely used in factor analysis, item response theory (a two-parameter normal ogive model) and genetic analyses (Neale and Cardon, 1992). In principle, the analysis involves numerical integration over as many non-missing items as exist in the response pattern of the subject. In practice, this method is computationally demanding (especially for the two-stem and 18-item DSM-III-R diagnostic criteria) so we used a marginal maximum likelihood approach (Bock, Gibbons and Muraki, 1988) implemented as a mixture distribution in Mx (Neale, Boker, Xie and Maes, 2002) using ordinal data analysis as described by Neale, Aggen, Kubarych, Foley and Kendler (in preparation) and Mehta, Neale and Flay (in press). The results would probably differ if the data were not MAR.

Factor analysis provides a method to investigate sources of covariation between different measures. It allows examination of how well the items in a scale each assess a latent trait. The equivalence of the factor model to the normal ogive item response theory (IRT) model has been noted by others (Takane and DeLeeuw, 1987). Another reason for the popularity of factor analysis is that factor scores can be computed and either correlated with measures of other constructs to help clarify the nature of the factors or used as predictor variables in multiple regression or dependent variables in ANOVA. There are several different methods of computing factor scores, however. Each of these methods has shortcomings and the choice of method can lead to different conclusions (Grice, 2001). Therefore, instead of using factor scores to predict other diagnoses for external validation, we explicitly entered variables representing lifetime major depression, phobias, alcohol dependence and panic disorder in further factor analyses. We have previously reported comorbidity between GAD and these disorders in women in terms of both tetrachoric correlations and odds ratios (Kendler, Walters, Neale, Kessler, Heath and Eaves, 1995). Tetrachoric correlations were 0.677 (GAD and major depression), 0.382 (GAD and phobias), 0.281

(GAD and alcoholism) and 0.484 (GAD and panic disorder).

Lifetime major depression was diagnosed using DSM-III-R criteria (American Psychiatric Association, 1987) and required five or more symptoms. Lifetime alcohol dependence was also based on DSM-III-R criteria and required three or more of nine symptoms for at least 28 days. Phobia was diagnosed using an adaptation of DSM-III criteria (American Psychiatric Association, 1980), which required the presence of one or more of 22 fears that the respondent recognized as unreasonable and that, in the judgment of the interviewer, objectively interfered with the respondent's life (Kendler, Myers, Prescott and Neale, 2001). The diagnosis of panic was based on DSM-III-R and required four symptoms that peak within 10 minutes and either four attacks within one month or that the patient worries about future attacks (American Psychiatric Association, 1987). We have shown that these approaches reflect the same continuum of liability as the fully syndromal disorders (Kendler, Gardner and Prescott, 2001). Although these variables are based on symptom criteria, not formal diagnoses, we adopt the shorthand of referring to them as 'diagnoses' below. All 'diagnoses' are lifetime.

Results

Endorsement frequency

The first two columns of Table 1 give the maximum likelihood estimates of the thresholds for the 18 DSM-III-R and 6 DSM-IV GAD symptoms in z-scores (column 1) and the probability of being above this threshold in a normal distribution (column 2). For example, with a z-score of 1.00, the threshold at which a subject would be expected to endorse the second stem is one standard deviation above the mean, so about 16% of subjects would be expected to endorse this item. The third column gives the observed endorsement rates among the 900 subjects who answered 'yes' to one or both of the stems. For every symptom, the endorsement rates of the 900 subjects who answered 'yes' to a stem question exceed or equal the proportion that would be expected to endorse the item in a normal population.

Thus, the general pattern is as expected: subjects endorsing either of the stems have liability higher than the population mean. These results substantially replicate the Marten et al. (1993) rates and verify that the symptoms with the highest endorsement frequencies were retained in DSM-IV. In only one case did a

symptom that was dropped from DSM-IV have an endorsement frequency (column 3) that exceeded that of any of the symptoms retained: 'Was your stomach often upset, or did you have nausea or diarrhea?' (0.581) was higher than both 'Did your muscles often feel tense, sore or achy?' (0.567) and 'Were you so nervous you had trouble concentrating?' (0.529).

Factor analyses

Confirmatory factor analyses were undertaken for both one and two factors. For the DSM-III-R symptoms, the chi-square difference between the one- and two-factor models of 144.5 on 19 degrees of freedom was highly significant, rejecting the hypothesis that DSM-III-R GAD is a unidimensional construct. For the DSM-IV symptoms, the chi-square difference of 68 on 9 degrees of freedom was also highly significant, rejecting the hypothesis that DSM-IV GAD is a unidimensional construct.

A series of model-fitting analyses were performed using Mx (Neale et al., 2002) to address comorbidity with other psychiatric disorders. Initially, we separately factor analysed the GAD symptomatic criteria for DSM-III-R and DSM-IV along with the two stems. In the DSM-IV case, it was necessary for identification purposes to fix one loading on one of the factors. We decided to fix the loading of the first stem to zero on the second factor (which we did for both DSM-III-R and DSM-IV). We reasoned that the first stem, about feeling anxious or worried most of the time, was the most general indicator of GAD; fixing it to zero on one factor shows its relationships to the other factor. Next, we added diagnoses for lifetime major depression, phobias, alcohol dependence and panic as variables to the factor analyses, and performed a series of rotations to examine their relationships to the factors.

Table 2 combines the results of the factor analyses for DSM-III-R and DSM-IV, with and without the added variables for major depression, alcohol dependence, any phobia and panic. In the analysis with the added variables, the loadings for the stems and probes were fixed to the values obtained in our analyses to show the relationships of major depression, alcohol dependence, phobias and panic to these factors.

DSM-III-R criteria

Both stem items loaded highly on the first factor (0.75 and 0.86). Six probes had significant (>0.4) loadings

Table 1. Comparison of theoretical probability of being above threshold with observed endorsement rates for DSM-III-R and DSM-IV GAD symptoms

Item	Threshold (z score)	p > threshold	Obs 1 or more probe = 1 (N = 900)
Stems			
In the last year, have you had a time lasting at least 5 days, when you felt anxious, nervous or worried most of the time?	0.25	0.401	0.950
In the last year, have you had a time lasting at least 5 days, when your muscles felt tense, or you felt jumpy or shaky inside?	1.00	0.159	0.352
DSM-III-R			
Did you often tremble, twitch or feel shaky?	0.79	0.216	0.382
* Did your muscles often feel tense, sore or achy?	0.47	0.318	0.567
† Did you often feel physically restless – couldn't sit still?	-0.08	0.532	0.706
* Did you often tire easily?	0.17	0.431	0.610
Did you often feel short of breath?	1.35	0.088	0.151
Did your heart often pound or race?	0.76	0.225	0.366
Did you sweat a lot? Were your hands often cold and clammy?	0.94	0.174	0.276
Did your mouth often feel dry?	0.81	0.210	0.320
Did you often feel dizzy or lightheaded?	1.13	0.128	0.232
Did you often have flushes (hot flashes) or chills?	1.10	0.137	0.243
Was your stomach often upset, or did you have nausea or diarrhea?	0.08	0.466	0.581
Did you urinate more often than usual?	0.98	0.164	0.231
Did you have trouble swallowing, or get a lump in your throat?	0.93	0.177	0.246
† Did you often feel keyed up or on edge?	-0.79	0.784	0.903
Did sudden noises often startle you?	0.70	0.241	0.373
* Were you so nervous you had trouble concentrating?	0.31	0.352	0.529
* Did you often have trouble falling asleep?	-0.11	0.543	0.659
* Were you often irritable or especially impatient?	-0.29	0.613	0.773
DSM-IV			
Did your muscles often feel tense, sore or achy?	0.44	0.332	0.567
Did you often tire easily?	0.17	0.433	0.610
Were you so nervous you had trouble concentrating?	0.38	0.353	0.529
Did you often have trouble falling asleep?	-0.11	0.544	0.659
Were you often irritable or especially impatient?	-0.28	0.610	0.773
Did you feel restless or keyed up and on edge?	-0.96	0.831	0.946

* = Symptom retained in DSM-IV. † = Combined into one symptom in DSM-IV.

on both factors. These included cardiovascular symptoms: 'Did you often feel short of breath?' (0.46 and 0.48) and 'Did your heart often pound or race?' (0.47 and 0.59). Also loading on both factors were 'Did you often tremble, twitch or feel shaky?' (0.61 and 0.50); 'Did you often feel dizzy or lightheaded?' (0.55 and 0.43); 'Did you often have flushes (hot flashes) or chills?' (0.52 and 0.53); and 'Did sudden noises often startle you?' (0.47 and 0.44).

Factor 1 also has substantial loadings for 'Did your muscles often feel tense, sore or achy?' (0.83); 'Did you often feel physically restless – couldn't sit still?' (0.52); 'Did you often tire easily?' (0.60), 'Did you often feel keyed up or on edge?' (0.57), 'Were you so nervous you had trouble concentrating?' (0.59) and 'Were you often irritable or especially impatient?' (0.55). This factor contains items that overlap with symptoms of Major Depressive Disorder (MDD), such as tiring

Table 2. Factor analysis of stems and DSM-III-R and IV probes. Estimates for depression, alcohol dependence, phobia and panic attack

	F1	F2
DSM-III-R		
(Stem) In the last year, have you had a time lasting at least 5 days when you felt anxious, nervous or worried most of the time?	0.75	=0.00
(Stem) In the last year, have you had a time lasting at least 5 days when your muscles felt tense, or you felt jumpy or shaky inside?	0.86	0.13
Did you often tremble, twitch or feel shaky?	0.61	0.50
* Did your muscles often feel tense, sore or achy?	0.83	0.12
† Did you often feel physically restless – couldn't sit still?	0.52	0.36
* Did you often tire easily?	0.60	0.11
Did you often feel short of breath?	0.46	0.48
Did your heart often pound or race?	0.47	0.59
Did you sweat a lot? Were your hands often cold and clammy?	0.37	0.63
Did your mouth often feel dry?	0.35	0.59
Did you often feel dizzy or lightheaded?	0.55	0.43
Did you often have flushes (hot flashes) or chills?	0.52	0.53
Was your stomach often upset, or did you have nausea or diarrhea?	0.34	0.27
Did you urinate more often than usual?	0.27	0.40
Did you have trouble swallowing, or get a lump in your throat?	0.21	0.53
† Did you often feel keyed up or on edge?	0.57	0.19
Did sudden noises often startle you?	0.47	0.44
* Were you so nervous you had trouble concentrating?	0.59	0.25
* Did you often have trouble falling asleep?	0.35	0.25
* Were you often irritable or especially impatient?	0.55	0.21
Major depression	0.49	0.24
Alcohol dependence	0.33	0.26
Any phobia	0.34	0.30
Panic	0.41	0.48
DSM-IV		
(Stem) In the last year, have you had a time lasting at least 5 days when you felt anxious, nervous or worried most of the time?	0.86	0.00
(Stem) In the last year, have you had a time lasting at least 5 days when your muscles felt tense, or you felt jumpy or shaky inside?	0.74	0.41
Did your muscles often feel tense, sore or achy?	0.70	0.71
Did you often tire easily?	0.59	0.21
Were you so nervous you had trouble concentrating?	0.82	0.05
Did you often have trouble falling asleep?	0.56	0.03
Were you often irritable or especially impatient?	0.67	0.11
Did you feel restless or keyed up and on edge?	0.83	0.08
Major depression	0.59	0.08
Alcohol dependence	0.45	-0.04
Any phobia	0.43	0.05
Panic	0.55	0.01

N = 2163. Loadings > 0.4 in boldface. * = Symptom retained in DSM-IV. † = Combined into one symptom in DSM-IV. F1 = Factor 1, F2 = Factor 2

easily and difficulty concentrating. When the diagnoses were added as variables, major depression loaded on this factor (0.49) and not significantly on factor 2.

Factor 2 has high loadings for 'Did you sweat a lot? Or did your hands often feel cold and clammy?' (0.63); 'Did your mouth often feel dry?' (0.59); 'Did you have

trouble swallowing or a lump in your throat?" (0.53); and 'Did you urinate more often than usual?' (0.40). This factor clearly contains the DSM-III-R autonomic cluster items: sweating a lot, dry mouth, and trouble swallowing. When the diagnoses were added, panic disorder loaded higher on this factor (0.48) than on factor 1 (0.41).

Three items, 'Was your stomach often upset, or did you have nausea or diarrhea?', 'Did you urinate more often than usual?' and 'Did you often have trouble falling asleep?' did not have significant loadings on either factor. Alcohol dependence and phobias did not load significantly on either factor.

DSM-IV criteria

For the DSM-IV criteria, factor 1 contains both stems (0.86 and 0.74). The second stem also reaches significance on factor 2 (0.41). One item loaded on both factors: 'Did your muscles often feel tense, sore or achy?' (0.70 and 0.71). Factor 1 has substantial positive loadings for all the probes. The highest is for the item that combined two DSM-III-R criteria, 'Did you feel restless or keyed up and on edge?' (0.83). The next highest loading was for 'Were you so nervous you had trouble concentrating?' (0.82). The other loadings were: 'Did you often tire easily?' (0.59), 'Did you often have trouble falling asleep?' (0.56), and 'Were you often irritable or especially impatient?' (0.67). The fact that muscle tension loads on the same factor as the other anxiety symptoms and all four diagnoses is important for Jorman and Stöber's (1999) suggestion that emphasizing muscle tension and de-emphasizing difficulty in concentration may improve discriminant validity between GAD and major depression. This does not appear to be the case in these data. All of the DSM-IV GAD symptoms, including difficulty concentrating, and all four diagnoses load on the first factor. Difficulty concentrating has a higher loading than muscle tension.

Relationships of diagnoses to stems and probes

In Table 3 we show the relationship of each of the four diagnoses to the first factor by fixing its loading to zero on the second factor. For both the DSM-III-R and DSM-IV cases, all four diagnoses show similar patterns. Again, it does not appear from these data that discriminant validity between GAD and major depression would be improved by emphasizing muscle tension and de-emphasizing difficulty concentrating.

For DSM-III-R, muscle tension loaded on the first factor 0.80, 0.74, 0.71 and 0.63 for MDD, alcohol dependence, phobia and panic respectively; the corresponding loadings for difficulty concentrating were 0.64, 0.62, 0.61 and 0.57, respectively. For DSM-IV, the figures are even less distinguishable: muscle tension loads 0.79, 0.64, 0.78 and 0.72 for MDD, alcohol dependence, phobia and panic respectively. The loadings for difficulty concentrating were virtually indistinguishable: 0.82, 0.81, 0.82 and 0.82 for MDD, alcohol dependence, phobia and panic respectively.

Tests of non-normality of factor space

To assess whether the distribution of the factors deviated from a normal distribution, the Gaussian weights used for quadrature were replaced by free parameters, thereby imposing an unrestricted shape for the latent factor distributions (Henkelman, Kay and Bronskill, 1990). This procedure was applied to the two-factor solution only, since the one-factor solution did not fit the data. No significant improvement in fit was obtained when the latent factors were allowed to be non-normal (chi-squared = 9.01, d.f. = 18, $p = 0.96$).

Discussion

In this study we conducted the first analyses to take account of the stem-and-probe format of the structured clinical interview, which forms the basis of most psychiatric research. We also conducted the first factor analysis of DSM-III-R GAD symptoms in a population-based sample and the first joint factor analysis of DSM-IV GAD symptoms and stem items in the SCID. As in the clinical sample of Kenardy et al. (1992), the DSM-III-R symptoms are clearly not unidimensional, and in our sample autonomic symptoms loaded on a factor related to panic but not depression. The hypothesis that the DSM-IV GAD symptoms represent a unidimensional construct was also rejected. The differences between the one- and two-factor models for DSM-IV were moderate but the attempt to render GAD a unidimensional construct by reducing the number of symptoms has not been completely successful.

The decision to eliminate the autonomic symptom cluster from GAD in DSM-IV was based on reliability and frequency of endorsement considerations (Marten et al., 1993). Our findings support at least the endorsement frequency rationale for the reduction in symptom criteria between DSM-III-R and the choice

Table 3. Factor loadings of stems, probes and lifetime diagnoses of MDD, phobias, alcohol dependence and panic disorder rotated to fix diagnoses to zero on second factor

DSM-III-R	Fix MDD = 0		Fix Alc Dep = 0		Fix Phobia = 0		Fix Panic = 0	
	F1	F2	F1	F2	F1	F2	F1	F2
(Stem) In the last year, have you had a time lasting at least 5 days when you felt anxious, nervous or worried most of the time?	0.67	-0.33	0.59	-0.46	0.56	-0.50	0.48	-0.57
(Stem) In the last year, have you had a time lasting at least 5 days when your muscles felt tense, or you felt jumpy or shaky inside?	0.83	-0.26	0.76	-0.42	0.73	-0.47	0.65	-0.57
Did you often tremble, twitch or feel shaky?	0.77	0.18	0.79	0.02	0.79	-0.03	0.78	-0.15
* Did your muscles often feel tense, sore or achy?	0.80	-0.25	0.74	-0.41	0.71	-0.46	0.63	-0.56
† Did you often feel physically restless – couldn't sit still?	0.63	0.09	0.63	-0.04	0.63	-0.08	0.61	-0.17
* Did you often tire easily?	0.59	-0.16	0.54	-0.28	0.52	-0.31	0.47	-0.39
Did you often feel short of breath?	0.63	0.23	0.66	0.10	0.66	0.06	0.67	-0.04
Did your heart often pound or race?	0.69	0.32	0.74	0.18	0.75	0.13	0.76	0.02
Did you sweat a lot? Were your hands often cold and clammy?	0.60	0.40	0.67	0.27	0.69	0.23	0.71	0.12
Did your mouth often feel dry?	0.57	0.37	0.63	0.25	0.65	0.21	0.67	0.11
Did you often feel dizzy or lightheaded?	0.68	0.14	0.70	0.00	0.70	-0.04	0.68	-0.15
Did you often have flushes (hot flashes) or chills?	0.70	0.25	0.74	0.10	0.74	0.05	0.74	-0.06
Was your stomach often upset, or did you have nausea or diarrhea?	0.42	0.09	0.43	0.00	0.43	-0.03	0.43	-0.09
Did you urinate more often than usual?	0.42	0.24	0.45	0.15	0.46	0.12	0.48	0.05
Did you have trouble swallowing, or get a lump in your throat?	0.42	0.38	0.49	0.29	0.51	0.25	0.54	0.18
† Did you often feel keyed up or on edge?	0.60	-0.08	0.57	-0.20	0.55	-0.24	0.51	-0.31
Did sudden noises often startle you?	0.61	0.19	0.64	0.06	0.64	0.02	0.63	-0.08
* Were you so nervous you had trouble concentrating?	0.64	-0.03	0.62	-0.16	0.61	-0.20	0.57	-0.29
* Did you often have trouble falling asleep?	0.43	0.07	0.43	-0.02	0.43	-0.05	0.42	-0.11
* Were you often irritable or especially impatient?	0.59	-0.05	0.57	-0.17	0.55	-0.21	0.52	-0.29
MDD	0.55	0.00	0.54	-0.11	0.53	-0.15	0.50	-0.22
Alcohol dependence	0.41	0.08	0.42	0.00	0.42	-0.03	0.41	-0.09
Any phobia	0.43	0.12	0.45	0.03	0.45	0.00	0.44	-0.07
Panic	0.58	0.26	0.62	0.12	0.63	0.09	0.63	0.00

Table 3. (contd)

DSM-IV	Fix MDD = 0		Fix Alc Dep = 0		Fix Phobia = 0		Fix Panic = 0	
	F1	F2	F1	F2	F1	F2	F1	F2
In the last year, have you had a time lasting at least 5 days when you felt anxious, nervous or worried most of the time?	0.85	-0.12	0.85	0.07	0.85	-0.11	0.86	-0.02
In the last year, have you had a time lasting at least 5 days when your muscles felt tense, or you felt jumpy or shaky inside?	0.79	0.30	0.70	0.47	0.79	0.31	0.75	0.39
Did your muscles often feel tense, sore or achy?	0.79	0.61	0.64	0.77	0.78	0.62	0.72	0.69
Did you often tire easily?	0.62	0.13	0.57	0.26	0.61	0.13	0.60	0.20
Were you so nervous you had trouble concentrating?	0.82	-0.07	0.81	0.12	0.82	-0.05	0.82	0.03
Did you often have trouble falling asleep?	0.56	-0.04	0.55	0.08	0.56	-0.04	0.56	0.02
Were you often irritable or especially impatient?	0.68	0.02	0.66	0.17	0.68	0.03	0.67	0.10
Did you feel restless or keyed up and on edge?	0.83	-0.03	0.82	0.16	0.83	-0.02	0.83	0.06
MDD	0.59	=0.00	0.58	0.16	0.59	0.01	0.59	0.07
Alcohol dependence	0.44	-0.10	0.45	=0.00	0.44	-0.10	0.45	-0.05
Any phobia	0.43	-0.01	0.42	0.09	0.43	=0.00	0.43	0.04
Panic	0.54	-0.06	0.54	0.06	0.54	-0.05	0.55	=0.00

of symptom criteria. It is also important to consider discriminant validity. All four diagnoses loaded on the same factor in DSM-IV for these data, suggesting less discriminant validity. Autonomic symptoms often help to discriminate between anxiety and depressive disorders, where low arousal is often seen (Turvey et al., 1999). Alternatively, autonomic symptoms may be a subsyndrome of panic disorder, which includes shortness of breath and racing heart.

The justification for using endorsement rates as the basis for including or not including symptoms in the definition of a disorder depends on the impact of the symptoms. The main purpose is to constrain the number of hypotheses as to which treatments will be effective. If the subjects who endorse low frequency items are more ill than subjects who endorse high frequency items, elimination of infrequently endorsed symptoms would reduce the effectiveness of the scale to differentiate between subjects with high levels of GAD. It is also possible that the infrequently endorsed symptoms are associated with more non-Axis I (clinical syndrome) pathology. We are not aware of any studies addressing whether GAD patients endorsing the infrequently endorsed symptoms that were dropped in DSM-IV are more likely to be diagnosed with particular personality disorders (Axis II) or personality traits, general medical conditions (Axis III), psychosocial stressors (Axis IV) or to have poorer global functioning (Axis V).

We did not find any significant improvement in fit when the factors were specified as having an arbitrary density function, as opposed to the normal distribution assumed in standard factor analyses. This finding indicates that the population distribution of liability to develop GAD (as assessed by the items for DSM-IV criteria) is consistent with a normal distribution. This finding is to be expected if liability to GAD is influenced by a large number of independent factors, each of small effect (the central limit theorem). Under these circumstances, attempts to characterize GAD dimensions into binary diagnostic categories would be inefficient for the purposes of research. Establishment of a threshold of severity above which subjects are diagnosed as cases may be a more practical approach for clinical purposes. Exactly which dimension, or what combination of the two dimensions found here, should be used is a matter for further research. The lack of evidence for non-normality in the two factors is consistent with Ruscio et al.'s (2001) conclusion that

pathological worry as represented by GAD may be quantitatively rather than qualitatively different from normal worry.

Faced with a multidimensional construct, there would appear to be two solutions. One factor could be eliminated if it was not considered to reflect the disorder. In the case of GAD, this is what has been done with the factor containing autonomic symptoms in the revision from DSM-III-R to DSM-IV, but even the latter still shows evidence for more than one factor. Alternatively, the inherently heterogeneous construct could be retained, but only on the basis of clinical intuition that they go together. This could result in a construct that may prove useful for clinical purposes but could hamper etiological research. Investigations into the validity and reliability of DSM diagnoses of GAD would probably prove less productive than studies of the factors underlying the liability to the disorder. Furthermore, subjects with GAD could prove inherently heterogeneous, some with high liability for one dimension and low for the second, and others having the reverse pattern of symptoms. Such uncontrolled heterogeneity would probably impair the identification of risk factors and sequelae of GAD and make replication of findings from such studies less likely to occur.

Krueger's (1999) study of 10 common mental disorders found an externalizing factor and an internalizing factor that split into subfactors of anxious misery and fear. Our DSM-III-R results could be interpreted as capturing anxious misery on factor 1 and fear on factor 2. Krueger argues that comorbidity results from common, core psychopathological processes, is a signal with important implications for nosology rather than noise, and that therefore research should focus on the common factor rather than separate disorders. Our analyses reveal that GAD factor 1 is associated with other DSM diagnoses such as MDD and Phobia. This is consistent with the perspective of Zinbarg and Barlow (1996) who argue that the DSM specifies a hierarchical model of anxiety disorders in which certain disorders share some common features.

Limitations

This study was limited to anxiety symptoms in women and may not generalize to men. These findings also pertain to the associated symptom criteria, not the diagnosis of GAD as a whole. Our 'subjects who met the criteria for GAD' are merely subjects who

endorsed six or more symptoms – subjects who were above a certain arbitrary cutoff on a symptom sum score. We did not impose any time restrictions on the GAD symptoms. DSM-IV diagnosis requires that the symptoms occur together more days than not for at least six months. We also do not have data in this interview for DSM-IV criteria B (the person finds it difficult to control the worry) and E (the symptoms cause impairment in functioning) and do not deal with hierarchy – subjects were not excluded if they met the criteria for MD. The diagnoses for depression, alcohol dependence, phobia and panic are lifetime diagnoses. Our results may not generalize to a specific timeframe – say, within the last year. In our analyses, subjects were treated as independent despite belonging to twin pairs. This method is not likely to bias parameter estimates but may have slightly inflated model comparison chi-squared tests, which should therefore be regarded as slightly anticonservative. Correlated pairs of observations do not, however, usually have large impact.

Future directions

A programme of developing a set of measurement instruments coordinated with each other is a large undertaking, but may be needed. The stems have endorsement rates that are lower than some of the probes, which suggests either that administration of those probes might be unnecessary or that they might be better used as stems. The second stem, 'In the last year, have you had a time lasting at least 5 days, when your muscles felt tense, or you felt jumpy or shaky inside?' has a threshold over a standard deviation above the mean (1.03); only three probes have higher difficulties. Even the 'easy' stem, with a difficulty of 0.25, has a higher difficulty than four of the probes.

One possibility for improving reliability would be to use adaptive testing to match items with subjects in such a way that the standard error of measurement is reduced, which leads to improved reliability. Waller and Reise (1989) applied adaptive testing in clinical decision making and showed that individuals who were extreme on the absorption trait were identified with perfect accuracy using, on average, 25% of the available items.

Reducing the number of symptom criteria does not seem to have improved the discriminant validity of GAD. One method to assess discriminant validity would be to include measures of personality, such as

those from the five-factor model, in studies of comorbidity. Generalized Anxiety Disorder, the other anxiety disorders and depression are all characterized by high levels of neuroticism. In the variant of the five-factor model being refined by Costa and McCrae, the broad domain of neuroticism is composed of the lower-level facets of anxiousness, angry hostility, depressiveness, self-consciousness, impulsiveness and vulnerability (Costa and Widiger, 2002). It is possible that comorbidity among GAD, the other anxiety disorders and depression results from their shared high levels of global neuroticism. If this is the case then discriminant validity may be best assessed at the facet level, essentially controlling for the components common to all anxiety disorders.

Including measures of personality traits in future studies is also important in order to assess the value of the multi-axial system incorporated in the DSM since 1980. The theory behind the multi-axial system is that Axis I syndromes emerge from the patient's personality (Axis II) and psychosocial stressors (Axis IV). An anxiety disorder in a patient with a dependent personality undergoing a romantic breakup is different from an anxiety disorder in a patient with a negativistic (passive-aggressive) personality unable to get along with coworkers (Comer, 1992; Millon and Davis, 1996). Krueger, Caspi, Moffitt, Silva and McGee (1996) found relations between normal personality (assessed by Tellegen's Multidimensional Personality Questionnaire) and mental disorders were robust, that personality was particularly relevant to severe and comorbid cases, and that the relations were not affected by whether measurement of disorder was continuous or discrete. Since it is also possible that Axis I disorders covary not with the categorical personality disorders but with personality traits that exist within the normal population and cross diagnostic boundaries, it is important to measure normal personality traits, not just categorical personality disorders. These issues will loom large in discussions regarding DSM-V.

Acknowledgement

This work was supported by NIH grants MH-65322, MH-40828 and MH/AA/DA-49492. We acknowledge the contribution of the Virginia Twin Registry, now part of the Mid-Atlantic Twin Registry (MATR), to ascertainment of subjects for this study. The MATR, directed by Dr J Silberg and L Eaves, has received support from the National Institutes of Health, the Carman Trust and the WM Keck, John Templeton and Robert Wood Johnson Foundations.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3 edn. Washington DC: APA, 1980.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3 edn, revised. Washington DC: APA, 1987.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4 edn. Washington DC: APA, 1994.
- Bock RD, Gibbons R, Muraki E. Full-information item factor analysis. *Applied Psychological Measurement*, 1988; 12: 261–80.
- Clark LA. Dimensional approaches to personality disorder assessment and diagnosis. In CR Cloninger (ed.) *Personality and Psychopathology*. Washington DC: American Psychopathological Association, 1999.
- Comer RJ. *Abnormal Psychology*. New York: WH Freeman & Company, 1992.
- Costa PT, Widiger TA (eds) *Personality Disorders and the Five-Factor Model of Personality*. Washington DC: American Psychological Association, 2002.
- Enders CK. The impact of nonnormality on full information maximum likelihood estimation for structural equation models with missing data. *Psychological Methods* 2001; 6: 352–70.
- Grice JW. Computing and evaluating factor scores. *Psychological Methods* 2001; 6: 430–50.
- Henkelman RM, Kay I, Bronskill MJ. Receiver operator characteristic (ROC) analysis without truth. *Medical Decision Making* 1990; 10: 24–9.
- Hettema JM, Neale MC, Kendler KS. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *American Journal of Psychiatry* 2001; 158: 1568–78.
- Joormann J, Stöber J. Somatic symptoms of generalized anxiety disorder from the DSM-IV: associations with pathological worry and depression symptoms in a nonclinical sample. *Journal of Anxiety Disorders* 1999; 13: 491–503.
- Kenardy J, Evans L, Oei TPS. The latent structure of anxiety symptoms in anxiety disorders. *American Journal of Psychiatry* 1992; 149: 1058–61.
- Kendler KS, Gardner CO, Prescott CA. Panic syndromes in a population-based sample of male and female twins. *Psychological Medicine* 2001; 31: 989–1000.
- Kendler KS, Myers J, Prescott CA, Neale MC. The genetic epidemiology of irrational fears and phobias in men. *Archives of General Psychiatry* 2001; 58: 257–65.
- Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. Generalized anxiety disorder in women. *Archives of General Psychiatry* 1992; 49: 267–71.
- Kendler KS, Prescott CA. A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry* 1999; 56: 39–44.
- Kendler KS, Walters EE, Neale MC, Kessler RC, Heath AC, Eaves LJ. The structure of genetic and environmental risk factors for six major psychiatric disorders in women. *Archives of General Psychiatry* 1995; 52: 374–83.
- Krueger RF. The structure of common mental disorders. *Archives of General Psychiatry* 1999; 56: 921–6.
- Krueger RF, Caspi A, Moffitt TE, Silva PA, McGee R. Personality traits are differentially linked to mental disorders: a multitrait-multidiagnosis study of an adolescent birth cohort. *Journal of Abnormal Psychology* 1996; 105: 299–312.
- Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: Wiley & Sons, 1987.
- Loranger AW. Categorical approaches to assessment and diagnoses of personality disorders. In CR Cloninger (ed.) *Personality and Psychopathology*. Washington DC: American Psychopathological Association, 1999.
- Marten PA, Brown TA, Barlow DH, Borkovec TD, Shear MK, Lydiard RB. Evaluation of ratings comprising the associated symptom criterion of DSM-III-R generalized anxiety disorder. *Journal of Nervous and Mental Disease* 1993; 181: 676–82.
- Mehta PD, Neale MC, Flay BR. Squeezing interval change form ordinal panel data: latent growth curves with ordinal outcomes. *Psychological Methods* (in press).
- Merikangas KR, Zhang H, Avenevoli S, Acharyya S, Neuenschwander M, Angst J. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich cohort. *Archives of General Psychiatry* 2003; 60: 993–1000.
- Millon T, Davis RD. *Disorders of Personality: DSM-IV and Beyond*. New York: Wiley, 1996.
- Moras K, Borkovec TD, DiNardo PA, Rapee R, Riskind J, Barlow DH. Generalized anxiety disorder. In TA Widiger, AJ Frances, HA, Pincus, R Ross, MB First, W Davis, Kline M (eds) *DSM-IV Sourcebook, Volume 2*. Washington DC: American Psychiatric Association, 1996.
- Neale MC, Aggen SH, Kubarych TS, Foley DL, Kendler KS. Research design and methods for the analysis of structured clinical interview data. Richmond VA: Department of Psychiatry, Medical College of Virginia of Virginia Commonwealth University, in preparation.
- Neale MC, Boker SM, Xie G, Maes HH. *Mx: Statistical modeling*. Richmond VA: Department of Psychiatry, Medical College of Virginia of Virginia Commonwealth University, 2002.
- Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. New York: Kluwer Academic, 1992.
- Neale MC, Eaves LJ, Kendler KS. The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics* 1994; 24: 239–58.

- Pickles A, Angold A. Natural categories or fundamental dimensions: on carving nature at the joints and the rearticulation of psychopathology. *Development and Psychopathology* 2003; 15: 529–51.
- Ruscio AM, Borkovec TD, Ruscio J. A taxonomic investigation of the latent structure of worry. *Journal of Abnormal Psychology* 2001; 413–22.
- Spitzer RL, Williams JBW. *Structured Clinical Interview for DSM-III-R (SCID)*. New York: Biometrics Research Department, New York Psychiatric Institute, 1985.
- Takane T, De Leeuw J. On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika* 1987; 52: 393–408.
- Turvey CL, Stevens DE, Merikangas, KR. The validity of the associated symptom criteria for DSM-IV generalized anxiety disorder. *International Journal of Methods in Psychiatric Research* 1999; 8: 129–37.
- Waller NG, Meehl PE. *Multivariate taxometric procedures. Distinguishing types from continua*. Thousand Oaks CA: Sage, 1998.
- Waller NG, Reise SP. Computerized adaptive personality assessment: an illustration with the absorption scale. *Journal of Personality and Social Psychology* 1989; 57: 1051–8.
- Widiger TA, Frances AJ, Pincus HA, Ross R, First MB, Davis W, Kline M (eds). *DSM-IV Sourcebook, Volume 4*. Washington DC: American Psychiatric Association, 1998.
- Zinbarg RE, Barlow DH. Structure of anxiety and anxiety disorders: a hierarchical model. *Journal of Abnormal Psychology* 1996; 105: 181–93.

Correspondence: Thomas S. Kubarych, Department of Psychiatry, Virginia Commonwealth University, PO Box 980126, Richmond VA 23298-1026, USA.
Telephone (+1) 804 828 8135.
Fax (+1) 804 828 1471.
Email: tkubarych@vcu.edu.