International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 18(4): 189–203 (2009)
Published online 23 December 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/mpr.298

The structure of mental disorders re-examined: Is it developmentally stable and robust against additions?†

HANS-ULRICH WITTCHEN,[1,2] KATJA BEESDO-BAUM,[1] ANDREW T. GLOSTER,[1] MICHAEL HÖFLER,[1] JENS KLOTSCHE,[1] ROSELIND LIEB,[2,3] ANDRÉ BEAUDUCEL,[4] MARKUS BÜHNER[5] & RONALD C. KESSLER[6]

1 Institute of Clinical Psychology and Psychotherapy, Technische Universitaet Dresden, Dresden, Germany
2 Max Planck Institute of Psychiatry, Munich, Germany
3 Institute of Psychology, Epidemiology and Health Psychology, University of Basel, Switzerland
4 Department of Differential Psychology, Helmut Schmidt Universitaet, Hamburg, Germany
5 Department of Psychology, Ludwig Maximilians Universitaet München, Germany
6 Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

Key words
structure, confirmatory factor analysis, exploratory factor analysis, classification

Correspondence
Hans-Ulrich Wittchen, PhD, Institute of Clinical Psychology and Psychotherapy, Technische Universitaet Dresden, Chemnitzer Str. 46, 01187 Dresden, Germany.
Telephone (+49) 351-46336983
Fax (+49) 351-46336984
Email: wittchen@psychologie.tu-dresden.de

†This article was published online on December 23, 2009. An error was subsequently identified. This notice is included in the online and print versions to indicate that both have been corrected (29/12/09).
‡Correction made here after initial publication.

Received 2 June 2009; revised 14 August 2009; accepted 14 October 2009

Abstract

Previous research suggests that patterns of comorbidity of a limited number of anxiety, depressive, substance use and antisocial personality (ASP) disorders among adults are best reflected by a hierarchical three-factor structure with two correlated factors ('anxious-misery' and 'fear') that are summarized in a second-order 'internalizing' factor and one 'externalizing' factor. It has not been examined whether this structure is developmentally stable and robust against additions of more diagnoses. Using data from a prospective-longitudinal community study of adolescents and young adults we re-evaluate the three-factor model originally proposed by Krueger (*Archives of General* Psychiatry, 1999; 56, 921–926). Using confirmatory factor analysis with identical conventions as in Krueger's original work we found that the three factor model did not fit robustly across age or a wider range of diagnoses. Using explanatory factor analysis we examined alternative structures. We found various clinically meaningful patterns with good fit that go substantially beyond the original three-factor structure. However, again, there is little consistency in findings when different age groups or different diagnoses are considered. Our findings suggest that psychopathology cannot be reduced to any simple structure. *Copyright © 2009 American Psychiatric Association. This article is being co-published by the International Journal of Methods in Psychiatric Research and the American Psychiatric Association*‡.

Introduction

Comorbidity among mental disorders is the rule rather than the exception. This fact holds true for both clinical and community samples in adults (Beesdo-Baum *et al.*, 2009a; Jacobi *et al.*, 2004; Kessler *et al.*, 2005; Wittchen and Jacobi, 2005) and youth (Beesdo-Baum *et al.*, in press; Fergusson *et al.*, 1993; Newman *et al.*, 1996; Wittchen *et al.*, 2000). This ubiquitous observation has stimulated several theories based on various methodological approaches. Hitherto, the most influential empirical approach used to account for high levels of comorbidity is confirmatory factor analysis (CFA). Within this tradition, extant studies report largely consistent findings of a three-factor model consisting of a hierarchical three-factor structure consisting of 'anxious-misery', 'fear', and 'externalizing', where the first two factors can be summarized into a second-order factor labeled 'internalizing' (Krueger, 1999) (see figure 1 in Beesdo-Baum *et al.*, this issue). These results were modeled using lifetime and 12-months diagnoses of 10 commonly occurring disorders for persons aged 15–54 in the cross-sectional National Comorbity Survey (NCS; Kessler *et al.*, 1994). The 10 diagnoses examined fit onto the factors as follows: Major depressive episode, dysthymic disorder, and generalized anxiety disorder had high loadings on the 'internalizing anxious-misery' factor. Social phobia, simple phobia, agoraphobia, and panic disorder had high loadings on the 'internalizing fear' factor. Alcohol dependence, drug dependence and antisocial personality (ASP) disorder had high loadings on the 'externalizing' factor.

Replication of this proposed three-factor structure has been achieved among adults from various countries for both lifetime (Cox *et al.*, 2002; Watson, 2005) and 12-month (Cox *et al.*, 2002; Slade and Watson, 2006; Vollebergh *et al.*, 2001) diagnoses and in a meta-analysis (Krueger and Markon, 2006a) (for overview see table 1 in Beesdo-Baum *et al.*, this issue). In sum, these findings have prompted discussions with far-reaching implications for research on basic science, etiology, and diagnostic classification (Andrews *et al.*, 2009; Goldberg *et al.*, 2009; Krueger and Markon, 2006b; Watson, 2005; Watson *et al.*, 2008). Some arguments suggest that this model is not only phenotypically relevant, but may represent an important organizing devise for understanding common psychopathological processes.

Despite the overall consistency of results based on comorbidity patterns of adults, a range of observations and considerations have stimulated considerable concerns about the validity and utility of these findings. Besides a number of methodological concerns regarding the factor analytic approach (Kessler *et al.*, 2005; Wittchen *et al.*, 1999a) and the way the data were analyzed and interpreted (Wittchen *et al.*, 1999a), particular questions arise as to whether the three-factor solution is *developmentally stable*, i.e. whether it appears robust against variations in age or sample composition. To date, only two studies have examined the factor structure of common mental disorders among youth. The first study found that a two-factor internalizing-externalizing model best fit the data in an unselected birth cohort of 18 and 21 year olds (Krueger *et al.*, 1998). The number and type of disorders used in this study were identical to those in the original Krueger (1999) analysis with the exception that obsessive compulsive disorder was included instead of panic disorder. Of note, this study examined the model fits of a one-factor, two-factor and four-factor models, but not a three-factor model. The second study, using identical methods and conventions as Krueger (1999), replicated the three-factor structure in a sample of adolescents and young adults from the prospective-longitudinal Early Developmental Stages of Psychopathology (EDSP) study (Beesdo-Baum *et al.*, this issue). However, this study was unable to fit the higher order structure such that the 'internalizing' factor summarizing 'anxious misery' and 'fear' had to be omitted due to poor fit between the hypothesized structure and the data. Thus, it remains an open question as to whether or not the three-factor structure is stable across various age groups.

Several considerations suggest that different latent factor structures of comorbidity may indeed exist at different developmental stages. For example, differences between anxiety and depressive disorders in age-of-onset patterns might lead to different comorbidity patterns in different age groups. Age-of-onset for anxiety disorders is predominantly in childhood whereas depressive disorders and substance use disorder shows a later onset in adolescence or adulthood (Beesdo-Baum *et al.*, 2009b; Kessler *et al.*, 2005; Newman *et al.*, 1996; Wittchen *et al.*, 2000; Wittchen *et al.*, 1999b).

In addition to concerns over the developmental stability of findings, it remains unclear whether the model is limited to the 10 diagnoses utilized or if it is robust to additions/deletions of disorders. Particularly, it is unclear whether the model is *stabile when more diagnoses are included* (compare table 2 in Beesdo-Baum *et al.*, this issue). The only previous study using a considerably larger number and range of diagnoses (19) performed an exploratory factor analysis using data from the National Comorbidity Survey Replication (NCS-R) (Kessler *et al.*, 2005). After excluding the disorders associated with negative correlations (obsessive-compulsive disorder and separation anxiety disorder), this study found a two-factor

solution with high factor loadings on the first factor for 'internalizing' disorders (anxiety disorders, major depressive episode) and on the second factor for 'externalizing' disorders (conduct disorder, substance disorders). Of note, five disorders had factor loadings of 0.30 or higher on both factors (dysthymia, mania/hypomania, oppositional defiant disorder, attention deficit hyperactivity disorder, and intermittent explosive disorder), although all five had higher loadings on the internalizing than externalizing factor. Therefore, questions arise as to whether the three-factor solution, or even a two-factor internalizing-externalizing solution, appears to appropriately reflect the structure of a broad range of mental disorders.

In summary, several studies have replicated Krueger's three-factor structure. However, these replications are mainly based on analyses using similar conventions and methods in terms of number and types of diagnosis included. Variations in sample or diagnostic composition may challenge the robustness of this finding. Therefore, the purpose of the current paper is to examine the stability of the three-factor structure of mental disorders in a prospective-longitudinal community sample of adolescents and young adults up to age 34. Specifically, we investigate (1) the *developmental stability* by examining Krueger's (1999) three-factor structure in different age spans and (2) the *stability to diagnostic additions* by examining the effects of using a considerably broader scope of diagnoses than in the original Krueger analysis and it's replications. In addition, we explore potentially better fitting, alternative models by means of exploratory factor analyses (a) using a broad range of diagnoses and (b) with a particular emphasis on anxiety.

This work was prepared in the context of the American Psychiatric Association (APA) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) considerations for a new meta-structure of mental disorders (Andrews *et al.*, 2009; Goldberg *et al.*, 2009; Wittchen *et al.*, 2009) and the DSM-V Anxiety, OC Spectrum, Posttraumatic, and Dissociative Disorder Work Group, who commissioned these analyses. It represents the work of the authors for consideration by the work group. Recommendations provided in this paper should be considered preliminary at this time; they do not necessarily reflect the recommendations or decisions for DSM-V, as the DSM-V development process is still ongoing.

Methods

Study design

The EDSP study (Lieb *et al.*, 2000; Wittchen *et al.*, 1998b; Wittchen *et al.*, 1998c) is a prospective-longitudinal study designed to collect data on the prevalence, incidence, comorbidity, risk factors and course of mental and substance use disorders in a representative sample of originally $N = 3021$ adolescents and young adults aged 14–24 years at baseline assessment. The study includes follow-up surveys (T1/T2/T3), a family history component (T0/T2/T3) and direct assessments of parents (T1/T3). The baseline sample was drawn in 1994 from government registries and is representative for residents aged 14–24 in the greater Munich area. The baseline (T0) was conducted in 1994/1995 among 14–24 year olds (weighted mean = 19.6, standard deviation (SD) = 3.3) and the response rate (RR) was 70.8%. At T1 only respondents of the younger study cohort (age 14–17 at T0) were interviewed ($N = 1228$; RR = 88.0%; interval since T0: 1.2–2.1 years). A total of 2548 out of the original baseline sample completed T2 (RR = 84.3%; interval since T0: 2.8–4.1 years) and 2210 completed T3 from 2002 through 2005 (RR = 73.2%; interval since baseline: 7.3–10.6 years); the age range at '10 year follow-up' was 21 to 34 years (mean = 28.0, SD = 3.4).

All participants provided written informed consent (for respondents aged 18 years and younger parental consent was provided). The EDSP project and its family genetic supplement have been approved by the Ethics Committee of the Medical Faculty of the Technische Universitaet Dresden (No: EK-13811).

Two kinds of datasets were generated: A 12-month diagnoses data set from all four assessment waves and a person-year data set with variables indicating whether a proband had ever fulfilled the diagnostic criteria for the disorders up through age t, with t ranging from age one up to the individual age at the last completed assessment (maximum age = 34; see the section on samples).

Diagnostic assessment

Individuals were interviewed face-to-face by trained clinical (mostly psychologists) interviewers using the computer-assisted lifetime (T0) and interval (T1, T2, T3) versions of the Munich-Composite International Diagnostic Interview (DIA-X/M-CIDI) (Wittchen *et al.*, 1998a; Wittchen and Pfister, 1997) and its embedded assessment modules (Lieb *et al.*, 2000; Wittchen *et al.*, 1998b; Wittchen *et al.*, 1998c), providing information on lifetime and 12-month symptoms, syndromes and diagnoses of 48 mental disorders. Clinical reappraisal studies have documented good reliability and validity of the DSM-IV diagnoses derived by the M-CIDI DSM-IV algorithms (Reed *et al.*, 1998; Wittchen, 1994; Wittchen *et al.*, 1998a).

First, the same 10 diagnoses as examined in the NCS-work by Krueger (1999) are considered here for the

Table 1 Diagnostic conventions for the ten diagnoses in the original Krueger model

Diagnosis	Definition	12 month data total $N = 9007$ N (%w)	Person-year data total $N = 75613$ N (%w)
MDE	According to M-CIDI DSM-IV algorithm	543 (6.2)	6585 (9.3)
Dysthmia	According to M-CIDI DSM-IV algorithm; not excluding hypomanic and manic episodes (in accordance to Krueger, 1999).	194 (2.2)	1736 (2.3)
GAD	According to M-CIDI DSM-IV algorithm	60 (0.8)	1137 (1.8)
Social phobia	According to M-CIDI DSM-IV algorithm impairment criterion only applied if age ≥ 18 at the respective assessment	188 (2.1)	2801 (3.6)
Specific phobia	According to M-CIDI DSM-IV algorithm impairment criterion only applied if age ≥ 18 at the respective assessment	541 (5.6)	8243 (9.7)
Agoraphobia	With or without panic attack (in accordance to Krueger, Archives 1999)	138 (1.6)	1687 (2.3)
Panic disorder	According to M-CIDI DSM-IV algorithm; that is, panic disorder with or without agoraphobia	209 (2.4)	2561 (3.3)
Alcohol dependence	According to M-CIDI DSM-IV algorithm	417 (5.0)	2900 (4.1)
Any illicit drug use dependence	According to M-CIDI DSM-IV algorithm	142 (1.5)	1111 (1.4)
CD or APD	CD, in 1053 completers of T1 family assessment: defined according to M-CIDI DSM-IV algorithm. At T2: defined as presence of at least three out of 15 CD symptoms assessed at T2 ($N = 2548$). Since there are no recency data, it was assumed that everyone who fulfilled criteria at T1 or T2, respectively was a 12-month case as well. APD: defined according to DSM-IV criteria (not part of M-CIDI DSM-IV algorithm, assessed at T2); must have been at least 18 years old at T2 (not the case for 230 out of 2548) and must have reported the above CD symptoms. CD or APD available only in 2638 cases who completed either the T1 family assessment or T2.	360 (4.4)	5063 (7.2)

Note: MDE, major depressive episode; GAD, generalized anxiety disorder; CD, conduct disorder; APD, antisocial personality disorder.

analysis of developmental stability. Table 1 shows the diagnostic conventions for the diagnoses along with their frequencies in the 12-month and person-year data set.

In order to examine the stability against adding more diagnoses, 18 diagnoses were consecutively submitted to the Krueger model. They are summarized in Table 2 with their diagnostic conventions and frequencies in the person-year data set in which this analysis was conducted.

Samples and missing values

12-month diagnoses data set. Overall, there were $N = 9007$ cases (interviews conducted) at all four assessments: $n = 1091$ cases were age 14–15, $n = 3333$ age 16–20, $n = 2878$ age 21–25, $n = 1323$ age 26–30 and $n = 382$ age 31–34. ASP disorder information was only available in $n = 3601$ cases who completed either the T1 family assessment [1053

Table 2 Overview over the 18 diagnoses added to the 10 diagnoses covered in the Krueger model and the conventions used

Diagnosis	Definition	*N* (%w) *N* total = 75 613
Specific phobia subtypes		
Animal	According to M-CIDI DSM-IV algorithm	3237 (3.9)
Natural environment	According to M-CIDI DSM-IV algorithm	2022 (2.3)
Blood injection	According to M-CIDI DSM-IV algorithm	2926 (3.5)
Situational	According to M-CIDI DSM-IV algorithm	1450 (1.8)
Other	According to M-CIDI DSM-IV algorithm	368 (0.4)
Hypomanic episode (HME)	According to M-CIDI DSM-IV algorithm	1331 (1.8)
Manic episode (MNE)	According to M-CIDI DSM-IV algorithm	940 (1.3)
Separation anxiety	According to M-CIDI DSM-IV algorithm in those who completed T1 (*N* = 1228 probands) assessment; *At* T0*:* defined as presence of at least three out of eight separation anxiety symptoms plus indication for distress or impairment derived by using items from the Retrospective Self-Report of Inhibition Scale. In those who did not complete the T1 assessment, no age of onset information was available; missing onsets were replaced with 11. Separation anxiety was defined as present if criteria at T0 or T1 were met.	1815 (2.4)
OCD	According to M-CIDI DSM-IV algorithm	657 (0.9)
PTSD	According to M-CIDI DSM-IV algorithm	797 (1.2)
Pain disorder	According to M-CIDI DSM-IV algorithm	3861 (4.9)
Hypochondriasis	According to M-CIDI DSM-IV algorithm	9 (0.01)
SSI 4/6	According to M-CIDI DSM-IV algorithm	1385 (1.9)
Psychotic disorder	Possible psychotic disorder (indicates possible presence of the following: schizophrenia, schizophreniform disorder, brief psychotic disorder, schizoaffective disorder), according to M-CIDI DSM-IV algorithm	107 (0.1)
Any eating disorder	According to M-CIDI DSM-IV algorithm	1336 (2.1)
ADHD	According to M-CIDI DSM-IV algorithm, only assessed in those who completed the T1-Family assessment	843 (3.7)[1]
ODD	According to M-CIDI DSM-IV algorithm, only assessed in those who completed the T1-family assessment	344 (1.6)[1]
Tics	Single item assessed in those who completed the T1-family assessment	203 (0.8)[1]
Elimination disorder	Enuresis or encopresis in those who completed the T1-family assessment, single items containing DSM-IV frequency and age criteria	858 (3.6)[1]

[1] Percentages refer to the subsample of those who completed the T1 family assessment (*N* = 23 438).

respondents with parental information on conduct disorder (CD)] or the T2 assessment ($N = 2548$) where both CD and antisocial personality disorder (APD) information was collected (see the section on statistical analysis later for handling missing values). For some analyses the N was smaller due to diagnoses not assessed in every assessment (see the section on statistical methods).

Person-year data set. A person-year data file was build with variables indicating whether a proband had ever fulfilled the diagnostic criteria for the respective disorder through age t, with t ranging from age one up to the individual age at the last completed assessment (maximum age = 34). For instance, a proband aged 20 at the last completed assessment contributes to the data set 20 observations. In cases where the reported age of onset for a particular disorder varied for a given respondent over time, the minimum reported age of onset was used. The dataset had a total of $n = 75\,613$ observations, $n = 39\,273$ in the age span 1–13, $n = 11\,954$ in the age span 14–17, $n = 10\,875$ in the age span 18–21, and $n = 13\,511$ in the age span 22–34. For some analyses the N was smaller due to diagnoses not assessed in every assessment or missing age of onset information in specific diagnoses [see the section on statistical methods and Beesdo-Baum *et al.* (this issue) for the definition of age of onset of ASP].

Statistical analyses

Software

All analyses [CFAs and exploratory factor analysis (EFAs)] were carried out in Mplus, Version 5. Krueger (1999) used PRELIS and LISREL for CFAs and seemingly cases with missing data were excluded. For every data set we ran three analyses dealing differently with missing values: (a) missing values were dealt with by using the full information maximum likelihood method based on the missing at random assumption; that is, whether values are missing is assumed to depend only on the values of other variables in the model, not on other variables; (b) complete case analysis; the model is only fitted among the observations with complete data; (c) missing values for ASP disorder were replaced with zeros. Missing values due to missing age of onset information of any diagnoses (in the person-year data set) were not coded as zero because they are known to be cases. These cases were omitted instead.

To ensure that the results were robust against choosing (a), (b) or (c) and since Mplus does not calculate the most important fit index the standardized root mean square residual (see section on model fit) in case (a) (because the therefore necessary residuals are missing for some cases) all three analyses were run in the replication part. The extensions analyses were restricted on method (c) because fit and factor loading results were almost identical in the replication part.

Weighted data

At T0 the probands had been sampled with different weights according to age (Wittchen *et al.*, 1998b; Wittchen *et al.*, 1998c). Therefore, in all analyses sampling weights were used to adjust the sample at T0 to the source population (Greater Munich in 1994/1995) with regard to age, sex and geographic location.

Model fit

As in Krueger (1999) the standardized root mean square residual (RMS) (usually abbreviated as SRMR; but we use the abbreviation 'RMS' as Krueger to avoid confusion) was chosen as the main index of model fit. RMS = 0 indicates perfect fit; RMS = 0.05 is considered good fit, and RMS = 0.08 adequate fit. RMS has been shown to be sensitive towards model misspecification and less sensitive than other global fit measures towards distribution and sample size in badly fitting covariance structure models (Hu and Bentler, 1998). We also calculated the CFI (Comparative Fit Index) and TLI (Tucker–Lewis Index; Relative Non-centrality Index) measures of model fit to confirm that the selection of the best-fitting model based on RMS was not sensitive towards the particular measure of fit used. Although Hu and Bentler (1998) recommend a cut-off value about 0.95 for the CFI, Sivo *et al.* (2006) found that the optimal cut-off value for the CFI decreases with sample size. Since sample size is very large in the present study, CFI values close to 0.90 are regarded as indicating acceptable model fit. TLI might become greater than one. Finally, the chi-squared goodness of fit test was calculated (irrespective of the limited value in very large samples when almost every model is rejected).

In all models that were assessed the loading between the second-order latent factor 'internalizing' and 'anxious misery' was predicted to be one or higher – at least if the standard error was taken into account. Technically, this means that a negatively definite covariance matrix between the latent variables occurred and that Krueger's model was logically inconsistent with our data. As a result it was not possible to statistically separate 'internalizing' and 'anxious misery', these factors were identical in our data. Note that this loading was also found to be almost perfect (0.93) in Krueger's original publication (1999). Instead we fitted models with the three latent dimensions 'anxious misery', 'fear' and 'externalizing' without

imposing a second order factor and freely estimated correlations.

Confirmatory factor analysis (CFA)

Consistent with Krueger (1999), the analysed matrix in CFA was the tetrachoric correlation matrix of the disorders. Unlike in the Krueger analysis, the present data are clustered within persons (because of longitudinal data with four assessment waves or the person-year data structure respectively). Here, simple weighted least squares yielded non-positive definite covariance matrixes and therefore the method was not used. Instead, the weighted least squares estimator based on a diagonal weight matrix was applied (known to yield more robust results for clustered data; Beauducel and Herzberg, 2006). In methods (b) and (c) the modelling of the mean structure had to be suppressed so that Mplus calculated RMS.

Exploratory factor analysis (EFA)

EFA was conducted with principal factor analysis based on tetrachorical correlations. The number of factors was determined by two means. First, the frequently used scree test was utilized. This test, however, is rather subjective and leads therefore to less clear results (Fabrigar *et al.*, 1999; Gorsuch, 1983). Therefore, parallel analyses were also conducted with the benefit of objectivity and precision (Beauducel, 2001) but tending to underestimate the number of factors if the first eigenvalue is high (Turner, 1998). With parallel analyses 100 data sets were simulated each with the same sample size as the sample in which the model is to be fitted. These data sets contain binary variables with success probabilities equal to the relative frequencies of the disorders in the respective EDSP data set. This procedure simulates drawing samples from a population in which these disorders are independent. The eigenvalues were averaged across the 100 replications and compared to the observed eigenvalues to determine the number of factors. With this approach one extracts as many factors as there are observed eigenvalues which are larger than the respective simulated eigenvalue.

Both methods were integrated when determining the number of extracted factors taking into account the fact that overextraction is less harmful than underextraction because with underextraction necessary information is lost (Fabrigar *et al.*, 1999). The scree test and parallel analysis results were sent to the coauthor André Beauducel who had no substantive interest in choosing a specific number of factors and was blind to the model results (factor loading, correlation and fit) associated with the different number of chosen factors. André Beauducel then decided how many factors to extract. This ensured that the uncertainties in determining these numbers was not misused for choosing a particular model substantively favored by the authors. When the number of dimensions was not clear as indicated by André Beauducel different solutions were considered comparing the factor loadings of different models with regard to substantive interpretation. If model fit was poor (RMS < 0.08) the solution with one additional factor was additionally considered.

The factors were rotated with oblique rotations using a tuning parameter of zero (quartimin rotation; see

Table 3 Blocks of diagnoses and associated factors added to Krueger's model in consecutive order

	To be added to
(a)	
HME	Anxious misery/externalizing
MNE	Anxious misery/externalizing
Separation anxiety	Separation anxiety/Fear
OCD	Anxious misery
PTSD	Anxious misery
Specific phobia subtypes[1]	Fear
(b)	
Pain disorder	Fear
Hypochondriasis	Fear
SSI4/6	Fear/anxious misery
(c)	
Psychotic disorder	Fear/anxious misery/externalizing
(d)	
Any eating disorder	Fear/anxious misery/externalizing
(e)[2]	
ADHD	Externalizing
ODD	Externalizing
Tics	Anxious misery/additional factor
Elimination disorder	Anxious misery/additional factor

[1] 'Any specific phobia' was omitted when adding the individual specific phobias. [2] Childhood disorders could only be added to the model in the subsample of those in which the T1-Family assessment was completed (N = 1053 out of 3021 individuals). In the associated analyses, hypochondria had to be omitted because there were no cases.

Harman, 1976). Since RMS is available in the EFA of Mplus the complete available information could be used without coding missings as zeros. As such, missings were treated with the full information maximum likelihood method in Mplus.

Specific analyses

(i) Developmental stability: CFAs as described earlier were used to assess whether the fit and factor loadings and correlations between them differ substantially over age (14–15, 16–20, 21–25, 26–30, 31–34 in the 12-month diagnoses data set, 1–13, 14–17, 18–21, 22–34 in the person-year data set).

(ii) Stability when adding more diagnoses: CFAs as described earlier were performed using the person-year data file with up to $N = 74\,634$ observations (missings occurred due to missing age of onset information). According to substantive considerations, the added diagnoses were grouped into blocks and variables from the same block were added together in consecutive order (see Table 3). Disorders for which factor assignment a priory was unclear were assigned according to the following hierarchy of criteria: (1) Minimal value of the loading of 'internalizing' on 'anxious misery' (before omitting 'internalizing', highest similarity to Krueger's model); (2) minimal RMS value; (3) a combination of maximal CFI and TLI values and maximal factor loadings.

(iii) Better fitting alternative models: EFAs in the person-year file were used to search for a model that might be more stable considering specific age groups and adding diagnoses to Krueger's model. Two sets of EFAs were run in the person-year data file: (a) using the whole sample ($N = 75\,613$ cases) while not considering childhood disorders only assessed in the young cohort (aged 14–17 at T0); (b) in the sample of the T1 family assessment completers ($N = 23\,438$) in which additional childhood disorders were assessed. The low prevalent diagnosis of hypochondriasis (nine cases in the entire sample) had to be omitted because otherwise the rotation algorithm did not provide converge. Twenty-three diagnoses were considered altogether in the total sample and 27 in the T1 family assessment completers with full childhood assessments.

The EFAs were also repeated in exactly the same way in the whole sample but restricted on the anxiety diagnoses generalized anxiety disorder (GAD), social phobia, the five different subtypes of specific phobia, separation anxiety, agoraphobia, panic disorder, post-traumatic stress disorder (PTSD) and obsessive compulsive disorder (OCD).

Results

Developmental stability: How robust is the three-factor model when examined in different age groups?

Using identical methods and procedures as Krueger (1999), we examined whether the three-factor solution and model fits substantially differ across the age spans (1–13, 14–17, 18–21, 22–34 years). These ages are considered a proxy for developmental stages ranging from childhood through adolescence to early adulthood. As in Beesdo-Baum *et al.* (this issue) for the total sample, the factor 'internalizing' had to be omitted since the covariance matrix between the latent variables was not positive definite at least after taking the standard errors into account. We found that the model fit was appropriate for some age groups, namely 14–17 year olds (RMS = 0.076) and 18–21 year olds (RMS = 0.066), but does not fit in the younger or the older cohorts (RMS of at least 0.089) (compare Table 4 upper part). Model fit was particularly poor in the age span 1–13 (RMS = 0.138).

Also, we repeated the CFA in specific age groups in the 12 months diagnoses data set (14–15, 16–20, 21–25, 26–30, 31–34). Here, after omitting 'internalizing' for the same reason, model fit was only satisfactory in the age group 21–25 (RMS = 0.071). In the other age groups RMS equaled at least 0.088 (16–20) with poor fit in the age groups 14–15 (0.144), 26–30 (0.147) and 31–34 (0.178).

Stability when more diagnoses are included

Additional disorders were assigned to the three assumed factors in a stepwise procedure (compare Table 3). Again, in all analyses the factor 'internalizing' had to be omitted since the covariance matrix between the latent variables was not positive definite at least after taking the standard errors of the loadings into account. We found that the addition of anxiety diagnoses and HME and MNE to the Krueger model (1999) still yielded marginally satisfactory model fit (RMS = 0.082, compare lower part of Table 4). After adding the somatoform disorders the model fit however became unsatisfactory (RMS = 0.109). Adding psychotic disorder further worsened the fit to 0.114 and particularly adding childhood disorder resulted in poor fit (RMS = 0.146).

Explorations into better factor solutions

Given that the three-factor CFA solutions were neither robust against additions nor developmentally stable, we

Table 4 Summary of the CFA results in specific ages and after adding diagnoses

Analysis/sample	Age range	*N*	Which model could be fitted?	RMS[5]	RMS[6] ≤ 0.08
Developmental stability					
12 month diagnoses at T0, T1, T2 and T3	14–34	9 007	'Internalizing' omitted*	0.044	X
	14–15	1 091	'Internalizing' omitted*	0.144	
	16–20	3 333	'Internalizing' omitted*	0.088	
	21–25	2 878	'Internalizing' omitted*	0.071	X
	26–30	1 323	'Internalizing' omitted*	0.147	
	31–34	382	'Internalizing' omitted*	0.178	
			'Internalizing' omitted*		
Person-year data	1–34	74 634	'Internalizing' omitted*	0.058	X
	1–13	38 779	'Internalizing' omitted*	0.101	
	14–17	11 803	'Internalizing' omitted*	0.076	X
	18–21	10 730	'Internalizing' omitted*	0.066	X
	22–34	13 322	'Internalizing' omitted*	0.089	
Stability when adding more diagnoses					
Person-year data, subsequently added diagnoses:					
specific phobia subtypes, other anxiety disorders, HME and MNE[1]	1–34	74 389	'Internalizing' omitted*	0.082	
somatoform disorders[2]	1–34	74 330	'Internalizing' omitted*	0.109	
psychotic disorder to 'externalizing'	1–34	74 330	'Internalizing' omitted*	0.114	
any eating disorder to 'anxious-misery'	1–34	74 278	'Internalizing' omitted*	0.112	
repeated in the data set of 1053 persons with T1-family assessment[3]	1–28	23 046	'Internalizing' omitted*	0.122	
childhood disorders[4]	1–28	22 930	'Internalizing' omitted*	0.146	

[1] OCD and PTSD, HME and MNE were assigned to 'anxious misery', separation anxiety was added to 'fear'.
[2] Pain disorder and hypochondrias were added to 'fear', SSI 4/6 was added to 'anxious misery'.
[3] ADHD and ODD where added to 'externalizing', tics and elimination disorder were added to 'anxious misery'.
[4] Only among those where information on childhood disorders is present.
[5] RMS = standardized root mean squared residual.
[6] RMS ≤ 0.08.

Int. J. Methods Psychiatr. Res. 18(4): 189–203 (2009). DOI: 10.1002/mpr

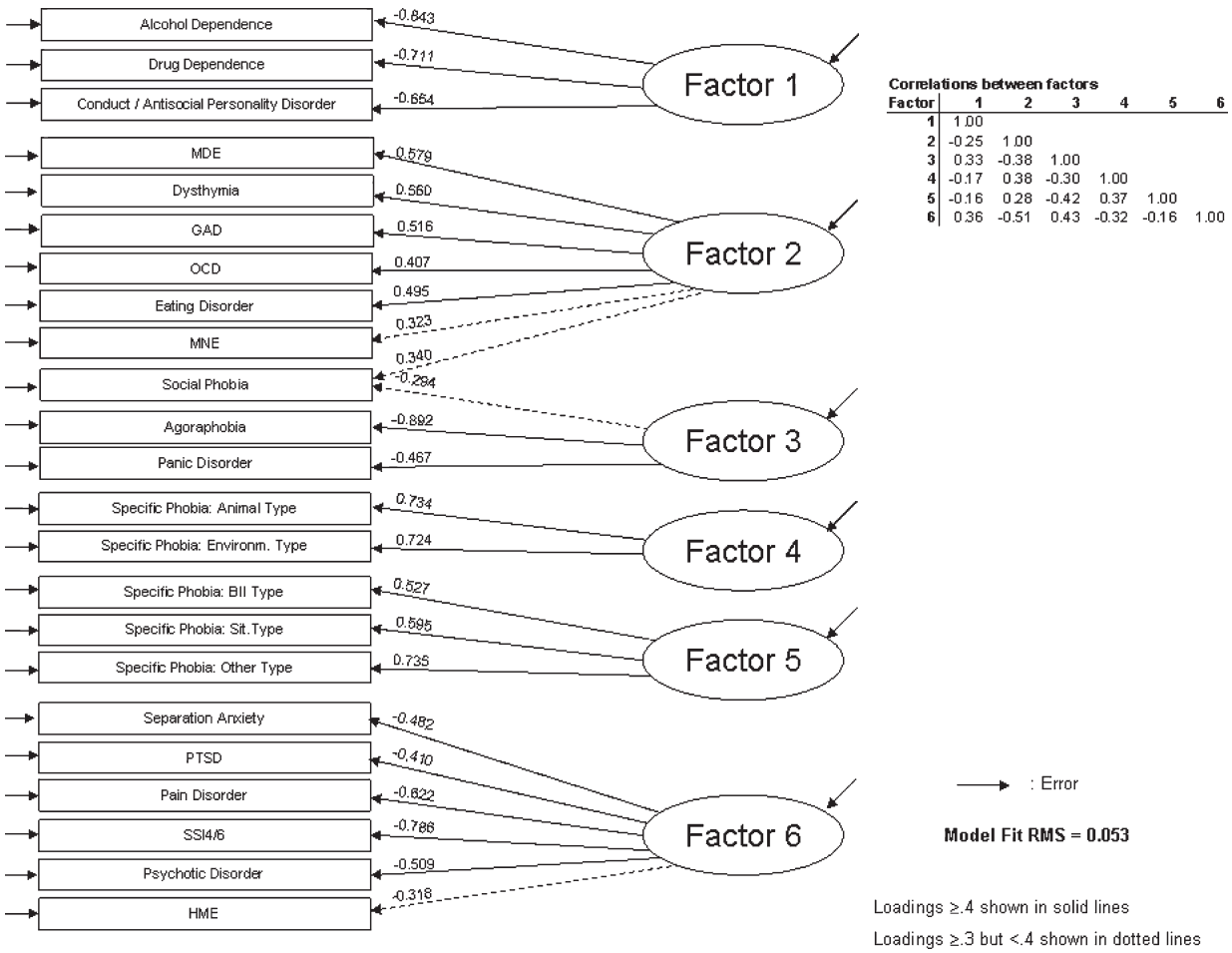

Figure 1 Result of the exploratory factor analysis (EFA) including anxiety, mood and other disorders (EDSP, total sample, six factors).

also explored whether other meaningful and stable factor models could be identified when using EFAs. Separate EFAs were conducted in the total person year data set and the age spans above after including the disorders listed in Table 2 (childhood disorders only assessed in the set of analyses in the young cohort). The scree test combined with parallel analyses suggested very different numbers of factors across age spans, with up to eight necessary factors. Also, the factors did not simply add up to subfactors. Therefore, it appears that there is no one single model that would apply to all different age spans examined.

An example of this exploration is presented here in Figure 1, namely the six factor solution from the total sample based on EFA with good fit values. Partly consistent with the three-factor model, this model describes a first 'externalizing' factor (alcohol and drug dependence, conduct/APD), a second factor that resembles the 'anxious-misery' factor [major depressive episode (MDE), dysthymia (DYS), and GAD), however with the addition of OCD and eating disorders. Krueger's (1999) 'fear factor' is however reflected by three factors: namely panic/agoraphobia, specific phobias (animal and environmental) and specific phobias (blood-injury, situational and other type). It is noteworthy that social phobia falls somehow in between factor two (anxious-misery) and three (panic and agoraphobia). Additionally, we found a sixth factor that describes a clinically more heterogeneous group of disorders, namely separation anxiety, PTSD, pain and somatoform disorders and psychotic disorder.

Table 5 summarizes the results across age spans. Whereas five factors were found in the age span 1–13 with MDE and dysthymia clustering with specific phobia, only three factors were found in the age span 14–17 with MDE and dysthymia now building a factor together with

Table 5 EFAs with added diagnoses across age spans

Age span		
1–13	Number of factors	RMS
	5	0.066
	Factors	*Included diagnoses*
	1	MDE, dysthymia, specific phobia blood injection, agoraphobia (loads negatively)
	2	GAD, HME, pain disorder, SSI 4/6
	3	Specific phobias animal and natural environment, MNE, OCD, any eating disorder
	4	Specific phobia animal (negative loading), PTSD
	5	Specific phobias blood injection, situational and other, agoraphobia, panic disorder, separation anxiety
	Unassigned diagnoses:[1]	Social phobia, ASP
	Omitted because too few cases:	Alcohol dependence, Illicit drug dependence
14–17	Number of factors	RMS
	3	0.066
	Factors	*Included diagnoses*
	1	Specific phobias situational and other (negative loadings), alcohol dependence, illicit drug dependence, ASP
	2	GAD, social phobia, specific phobias nat. environment, situational and other, agoraphobia, panic disorder, MNE, separation anxiety, OCD, PTSD, pain disorder, SSI 4/6
	3	MDE, dysthymia, psychotic disorder
	Unassigned diagnoses:[1]	Specific phobias animal and blood injection, any eating disorder
18–21	Number of factors	RMS
	6	0.057
	Factors	*Included diagnoses*
	1	MDE, dystymia, GAD, social phobia, OCD
	2	Psychotic disorder
	3	Specific phobia animal type
	4	Specific phobias situational and other, agoraphobia, panic disorder
	5	Alcohol dependence, illicit drug dependence, ASP
	6	SSI 4/6
	Unassigned diagnoses:[1]	Specific phobias nat. environment and blood injection, HME, MNE, separation anxiety, PTSD, pain disorder, any eating disorder
22–34	Number of factors	RMS
	7[2]	0.057
	Factors	*Included diagnoses*
	1	MDE, MNE
	2	Separation anxiety, SSI 4/6, psychotic disorder
	3	GAD, social phobia
	4	Specific phobias animal and nat. environment
	5	Alcohol dependence, illicit drug dependence, ASP
	6	Specific phobia blood injection, agoraphobia
	7	Specific phobia other, panic disorder
	Unassigned diagnoses:[1]	Dysthymia, HME, OCD, PTSD, pain disorder, any eating disorder

[1] No loading ≥ 0.4.

[2] It was unclear whether six or seven factors should be extracted. Here, we present the results for seven factors in order not to overlook an important dimension.

Int. J. Methods Psychiatr. Res. 18(4): 189–203 (2009). DOI: 10.1002/mpr

psychotic disorder. In the six factors required for the age span 18–21 MDE and dysthymia cluster together with GAD, social phobia and OCD, whereas in the seven factors model for age span 22–34 MDE and dysthymia fall into different factors with MDE clustering with MNE, dysthymia could not be clearly assigned to any factor (loadings < 0.4). Likewise, other disorders like GAD, social phobia and PTSD cluster with different other diagnoses across age spans.

This picture becomes even becomes more heterogeneous across age if the childhood disorders are added in the younger cohort with up to eight factors. For example, among the youngest cohort subjects a four-factor solution emerged: The first factor loads high on MDE (0.86), dysthmia (0.51), OCD (0.50) and eating disorders (0.48), the second on some specific phobias, the third on GAD, attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD) and elimination disorders, and the fourth on panic, agoraphobia, separation anxiety and some phobias (results available on request).

Nevertheless, the factor solutions may offer a range of clinically meaningful groupings – with some consistency across factor solutions: fairly consistently, an externalizing-like factor appeared (i.e. substance dependence, ASP), the frequent emergence of a panic/agoraphobia factor, a phobia factor (specific subtypes of phobia), and less consistently and depending on age group considered, a psychotic factor, an ADHD/ODD/CD factor, etc. It is also noteworthy, that some disorders such as OCD, psychotic, hypomania/mania did not consistently reveal particularly high loading on one single factor but rather displayed moderate loadings on several factors.

In a final step, we also subjected only the anxiety disorders (including separation anxiety disorder and specific phobia subtypes) to an EFA in order to explore whether a meaningful model could be identified within anxiety disorders. The scree plots and parallel analyses consistently indicated three factor solutions with acceptable model fits overall and by age group considered. Despite some differences in the loadings and the grouping of specific disorders there are common findings: (a) agoraphobia and panic disorder were always found in the same factor; (b) GAD and social phobia were always found in the same factor; (c) in all three factor solutions from age 14 on (and in the total sample) GAD, social phobia, PTSD, agoraphobia, panic disorder, OCD and separation anxiety where assigned, with some variations, to the same factor; (d) the specific phobia subtypes animal and natural environment are always assigned to the same factor; in the age spans 14 to 17 and 18 to 21 they form a factor on their own, in the age span 1 to 13 they cluster with GAD and social phobia, in the age span 22 to 34 with the situational subtype; (e) the blood injection injury specific phobia subtype could not be clearly assigned to one single factor.

Discussion

Replications of Krueger's (1999) three-factor structure of mental disorders, with two correlated 'internalizing' factors and one 'externalizing' factor, were mainly based on analyses using similar conventions and methods in terms of number and types of diagnosis included (Beesdo-Baum *et al.*, this issue; Krueger and Markon, 2006a). The current paper examined the stability of this structure in a prospective-longitudinal community sample of adolescents and young adults up to age 34. The rationale behind these analyses is as follows: If we find that Krueger's model fails to apply robustly across age and diagnoses, psychopathology as a whole seemingly can not be reduced to this structure.

Using CFAs we found that neither was the second order 'internalizing' factor consistent with the EDSP data (no matter whether diagnoses were added or not or the total sample or specific age groups were assessed) nor did the three factor model fit robustly across age or a wider range of diagnoses. Using EFAs we additionally examined alternative structures. We found various clinically meaningful patterns with good fit, that go substantially beyond the original Krueger (1999) three-factor structure. However, as with regard to the Krueger structure, there is little consistency in findings when different age groups or different diagnoses are considered suggesting that psychopathology cannot be reduced to any simple structure. This applies even more if one takes into account that all the analyses are inherently based on very strong assumptions and data reduction which may lead to a heavy underestimation of the true complexity of psychopathology: First, in his original work Krueger (1999) only examined four models when deciding on the candidate dimensions and factor assignments for them. These models were chosen based on prior explorative analyses of the same data. The conceptual justification for picking those restricted models and not others is quite limited, especially in light of Krueger's reliance on CFAs and the fact that he extends his model beyond adolescence (Wittchen *et al.*, 1999a).

Second, data are reduced to *binary diagnoses* of the DSM system and are not based on the basic psychopathological level at which they should be based; i.e. the level of *symptoms* (Wittchen *et al.*, 1999a).

Third, data is reduced to the matrix of *tetrachorical correlations* among the disorders; i.e. one assumes that the

comorbidity between each pair of diagnoses was independent of all other disorders. In a model with k disorders, 2^k disorder profiles are possible, but only $\binom{k}{2}$ free parameters are considered in the matrix of tetrachorical correlations. Kraemer (1997, p. 1121) has argued that this assumption is likely to be violated for mental disorders. Kessler *et al.* (2005) have demonstrated that this assumption does not apply to the 12-month diagnoses data from the NCS-R: more respondents did not fulfill the criteria of any disorders and fewer respondents were found with high comorbidity than assumed under the assumption that there were only two-way associations. They considered $k = 19$ disorders with 524 288 possible configurations 433 of which were actually observed but only 171 parameters are considered in the matrix of pairwise correlations. In the present EDSP person-year data the 10 diagnoses from the Krueger model yielded 416 different configurations but the associated matrix of tetrachorical correlations has only 45 free parameters. Adding 18 diagnoses 1592 different configurations were found as compared to 378 free parameters here.

Moreover, Kraemer (1997) has shown that 'for associations better than random . . . and less than perfect' the tetrachoric correlation coefficient exceeds the kappa statistic, resulting in an overestimation of the association and, thus, maybe bias in the factor structure.

There are more subtle problems with using (confirmatory and explanatory) factor analysis for constructing models for the structure of comorbidity of mental disorders: First, they are based on the assumption of an underlying normal continuum behind binary diagnoses although the construction of DSM up to version IV-TR (i.e. the rationale behind the diagnoses used by Krueger) assumes clear thresholds between cases and non-cases. More importantly, classical test theory where factor analysis comes from assumes that each item has the same difficulty for all respondents; i.e. that the size of the factor loadings does not depend on the item difficulty (note that using CFA with categorical variables in Mplus corresponds with multidimensional item response theory via the thresholds of the categories). However, in the present study the diagnoses have different probabilities, so that these probabilities might have affected the covariances which are the basis for factor analysis.

Having noted the several basic restrictions the question arises whether there are any more appropriate methods and approaches available. Latent class analyses (LCAs) have been successfully applied for modeling PTSD symptoms (Chung and Breslow, 2008) and mental disorders in general (e.g. Kessler *et al.*, 2005). Although LCA takes the complete complexity of comorbidity into account, it classifies individuals, not variables (diagnoses). However, by taking all parameters into account LCA is even more unlikely than factor analysis to yield robust findings which can be repeated across age and disorders considered. Hybrid models extend LCAs by taking a latent continuum (e.g. severity of general psychopathology) into account often yielding a better model fit (Muthen, 2006; Muthen and Asparouhov, 2006). By combining LCA with factor analysis these models are in line with the intended modification of the DSM system in version V of integrating categorical and dimensional features of psychopathology (Regier *et al.*, 2009). Hybrid models have been successfully applied e.g. for specific topics like nicotine dependence criteria (Muthen and Asparouhov, 2006) and alcohol abuse and dependence criteria (Muthen, 2006). However, the problems of complexity and classifying individuals, not characteristics remain.

To conclude, overall we caution the use of such statistical models and explorations in general and the considerable degree of speculation surrounding clinical issues, i.e. the use of such data for clinical diagnostic and classificatory issues of mental disorders, in particular. These methods – if applied to the data currently available – are neither sufficiently robust nor clinically sensitive to reflect the complexity of mental disorders and to test assumptions about the underlying processes of psychopathology. This limitation is particularly relevant in light of the fact that the type of data needed for such explorations are simply not available up to now. However, even if study designs would be available that generate such data, allowing the use of such statistical methods without major restrictions, it seems unlikely that fairly simple and robust structural models will ever be derived, given the complexity of psychopathological features across the lifespan. To date, we must conclude that there appears not to be a stable and robust model that can restrict psychopathology to few dimensions. Thus, suggestions and implications based on simple structure models appear overly simplistic and not reflective of the 'true' complexity of psychopathological processes.

Acknowledgements

This work has been funded by the American Psychiatric Association (APA) Task Force for DSM-V. The data used come from the Early Developmental Stages of Psychopathology (EDSP) Study, funded by the German Federal Ministry of Education and Research (BMBF) project nos 01EB9405/6, 01EB 9901/6, EB01016200, 01EB0140, and 01EB0440. Part of the field work and analyses were also additionally supported by grants of the Deutsche Forschungsgemeinschaft (DFG) LA1148/1-1, WI2246/1-1, WI 709/7-1, and WI 709/8-1.

Principal investigators of the EDSP are Dr Hans-Ulrich Wittchen and Dr Roselind Lieb.

Core staff members of the EDSP group are: Dr Kirsten von Sydow, Dr Gabriele Lachner, Dr Axel Perkonigg, Dr Peter Schuster, Dipl.-Stat. Michael Höfler, Dipl.-Psych. Holger Sonntag, Dipl.-Psych. Tanja Brückl, Dipl.-Psych. Elzbieta Garczynski, Dr Barbara Isensee, Dipl.-Psych. Agnes Nocon, Dr Chris Nelson, Dipl.-Inf. Hildegard Pfister, Dr Victoria Reed, Dipl.-Soz. Barbara Spiegel, Dr Andrea Schreier, Dr Ursula Wunderlich, Dr Petra Zimmermann, Dr Katja Beesdo-Baum, Dr Antje Bittner, Dipl.-Psych. Silke Behrendt, and Dipl.-Psych. Susanne Knappe. Scientific advisors are Dr Jules Angst (Zurich), Dr Jürgen Margraf (Basel), Dr Günther Esser (Potsdam), Dr Kathleen Merikangas (NIMH, Bethesda), Dr Ron Kessler (Harvard, Boston) and Dr Jim van Os (Maastricht).

Declaration of interest statement

Dr Beesdo-Baum has received speaking honoraria from Pfizer and Eli Lilly and Company. Dr Wittchen has received speaking honoraria from Eli Lilly and Company, Pfizer and Norvartis and serves on advisory boards for Pfizer, Servier and Schering-Plough. Dr Höfler, Dr Gloster, Dipl.-Stat. Klotsche, Dr Lieb, Dr Beauducel, Dr Bühner and Dr Kessler have nothing to declare.

References

Andrews G., Goldberg D.P., Krueger R.F., Carpenter W.T., Hyman S.E., Sachdev P., Pine D.S. (2009) Exploring the feasibility of a meta-structure for DSM-V and ICD-11: Could it improve utility and validity? *Psychological Medicine*, **39**, 1993–2000.

Andrews G., Pine D.S., Hobbs M.J., Anderson T.M., Sunderland M. (in press b) Neurodevelopmental disorders: Cluster 2 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine*.

Beauducel A. (2001) Problems with parallel analysis in data sets with oblique simple structure. *Methods of Psychological Research Online*, **6**, 2.

Beauducel A., Herzberg P. (2006) On the performance of maximum likelihood versus means and variance adjusted weighted least square estimation in confirmatory factor analysis. *Structural Equation Modeling*, **13**, 186–203.

Beesdo-Baum K., Höfler M., Gloster A.T., Klotsche J., Lieb R., Beauducel A., Bühner M., Kessler R.C., Wittchen H.-U. (in this issue) The structure of common mental disorders: A replication study in a community sample of adolescents and young adults. *International Journal of Methods in Psychiatric Research*.

Beesdo-Baum K., Jacobi F., Hoyer J., Low N.C.P., Höfler M., Wittchen H.U. (2009a) Pain associated with specific anxiety and depressive disorders in a nationally representative population sample. *Social Psychiatry and Psychiatric Epidemiology*, DOI: 10.1007/s00127-009-0045-1

Beesdo-Baum K., Knappe S., Pine D.S. (2009b) Anxiety and anxiety disorders in children and adolescents: Developmental issues and implications for DSM-V. *Psychiatric Clinics of North America*, **32**, 483–524.

Beesdo-Baum K., Pine D.S., Lieb R., Wittchen H.U. (in press) Incidence and risk patterns of anxiety and depressive disorders and categorization of Generalized Anxiety Disorder. *Archives of General Psychiatry*.

Chung H., Breslow N. (2008) The latent structure of posttraumatic stress disorder: tests of invariance by gender and trauma type. *Psychological Medicine*, **38**, 563–573.

Cox B.J., Clara I.P., Enns M.W. (2002) Posttraumatic stress disorder and the structure of common mental disorders. *Depression and Anxiety*, **15**, 168–171.

Fabrigar L., Wegener D., MacCallum R., Strahan E. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, **4**, 272–299.

Fergusson D.M., Horwood L.J., Lynskey M.T. (1993) Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15 year olds. *Journal of the American Academy of Child and Adolescent Psychiatry*, **32**, 1127–1134.

Goldberg D.P., Krueger R.F., Andrews G., Hobbs M.J. (2009) Emotional disorders: Cluster 4 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine*, **39**, 2043–2059.

Gorsuch R. (1983) *Factor Analysis*, Lawrence Erlbaum.

Harman H. (1976) *Modern Factor Analysis*, University of Chicago Press.

Hu L., Bentler P.M. (1998) Fit indices in covariance structure analysis: Sensitivity to under-parameterized model misspecification. *Psychological Methods*, **3**, 424–453.

Jacobi F., Wittchen H.-U., Hölting C., Höfler M., Pfister H., Müller N., Lieb R. (2004) Prevalence, comorbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychological Medicine*, **34**, 597–611.

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.

Kessler R.C., McGonagle K.A., Zhao S., Nelson C.B., Hughes M., Eshleman S., Wittchen H.-U., Kendler K.S. (1994) Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry*, **51**, 8–19.

Kraemer H.C. (1997) What is the 'right' statistical measure of twin concordance (or diagnostic reliability and validity)? *Archives of General Psychiatry*, **54**, 1121–1124.

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

Krueger R.F., Caspi A., Moffitt T.E., Silva P.A. (1998) The structure and stability of common mental disorders (DSM-III-R): A longitudinal-epidemiological study. *Journal of Abnormal Psychology*, **107**, 216–227.

Krueger R.F., Markon K.E. (2006a) Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, **2**, 111–133.

Krueger R.F., Markon K.E. (2006b) Understanding psychopathology. Melding behavior genetics, personality, and quantitative psychology to develop an empirically based model. *Current Directions in Psychological Science*, **15**, 113–117.

Lieb R., Isensee B., von Sydow K., Wittchen H.-U. (2000) The early developmental stages of psychopathology study (EDSP): A methodological update. *European Addiction Research*, **6**, 170–182.

Muthen B. (2006) Should substance use disorders be considered as categorical or dimensional? *Addiction*, **101**, 6–16.

Muthen B., Asparouhov T. (2006) Item response mixture modelling: Application to tobacco dependence criteria. *Addiction Behaviors*, **31**, 1050–1066.

Newman D.L., Moffitt T.E., Caspi A., Magdol L., Silva P.A. (1996) Psychiatric disorder in a birth cohort of young adults: Prevalence, comorbidity, clinical significance, and new case incidence from age 11 to 21. *Journal of Consulting and Clinical Psychology*, **64**, 552–562.

Reed V., Gander F., Pfister H., Steiger A., Sonntag H., Trenkwalder C., Sonntag A., Hundt W., Wittchen H.-U. (1998) To what degree the Composite International Diagnostic Interview (CIDI) correctly identify DSM-IV disorders? Testing validity issues in a clinical sample. *International Journal of Methods in Psychiatric Research*, **7**, 142–155.

Regier D.A., Narrow W.E., Kuhl E.A., Kupfer D.J. (2009) The conceptual development of DSM-V. *American Journal of Psychiatry*, **166**, 645–650.

Sivo S.A., Fan X., Witta E.L., Willse J.T. (2006) The search for 'optimal' cutoff properties: Fit index criteria in structural equation modeling. *Journal of Experimental Education*, **74**, 267–288.

Slade T., Watson D. (2006) The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychological Medicine*, **36**, 1593–1600.

Turner N. (1998) The effect of common variance and structure patterns on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement*, **58**, 541–568.

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.

Watson D. (2005) Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, **114**, 522–536.

Watson D., O'Hara M.W., Scott S. (2008) Hierarchical structures of affect and psychopathology and their implications for the classification of emotional disorders. *Depression and Anxiety*, **25**, 282–288.

Wittchen H.-U. (1994) Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*, **28**, 57–84.

Wittchen H.-U., Beesdo-Baum K., Gloster A.T. (2009) A new meta-structure of mental disorders: A helpful step into the future or a harmful step back to the past? A commentary on the Andrews series papers. *Psychological Medicine*, **39**, 2083–2087.

Wittchen H.-U., Höfler M., Merikangas K.R. (1999a) Towards the identification of core psychopathological processes? *Archives of General Psychiatry*, **56**, 929–931.

Wittchen H.-U., Jacobi F. (2005) Size and burden of mental disorders in Europe: A critical review and appraisal of 27 studies. *European Neuropsychopharmacology*, **15**, 357–367.

Wittchen H.-U., Kessler R.C., Pfister H., Lieb R. (2000) Why do people with anxiety disorders become depressed? A prospective-longitudinal community study. *Acta Psychiatrica Scandinavica*, **102**(Suppl. 406), 14–23.

Wittchen H.-U., Lachner G., Wunderlich U., Pfister H. (1998a) Test–retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Social Psychiatry and Psychiatric Epidemiology*, **33**, 568–578.

Wittchen H.-U., Lieb R., Schuster P., Oldehinkel A.J. (1999b) When is onset? Investigations into early developmental stages of anxiety and depressive disorders. In: *Childhood Onset of 'Adult' Psychopathology, Clinical and Research Advances* (ed. Rapoport JL), pp. 259–302, American Psychiatric Press.

Wittchen H.-U., Nelson C.B., Lachner G. (1998b) Prevalence of mental disorders and psychosocial impairments in adolescents and young adults. *Psychological Medicine*, **28**, 109–126.

Wittchen H.-U., Perkonigg A., Lachner G., Nelson C.B. (1998c) Early developmental stages of psychopathology study (EDSP) – objectives and design. *European Addiction Research*, 1998c; **4**, 18–27.

Wittchen H.-U., Pfister H. (1997) *DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); Ergänzungsheft (DIA-X-Lifetime); Interviewheft Querschnittuntersuchung (DIA-X-12 Monate); Ergänzungsheft (DIA-X-12Monate); PC-Programm zur Durchführung des Interviews (Längs- und Querschnittuntersuchung); Auswertungsprogramm*, Swets & Zeitlinger.