International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 18(4): 204–220 (2009)
Published online 18 December 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/mpr.293

The structure of common mental disorders: A replication study in a community sample of adolescents and young adults

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

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 Universität München, Munich, Germany
6 Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

Key words
epidemiology, confirmatory factor analysis, anxiety disorder, depressive disorder, substance use disorder, antisocial personality, community study

Correspondence
Katja Beesdo-Baum, PhD, Institute of Clinical Psychology and Psychotherapy, Technische Universitaet Dresden, Chemnitzer Str. 46, 01187 Dresden, Germany.
Telephone (+49) 351-463-36989
Fax (+49) 351-463-36984
Email: beesdo@psychologie.tu-dresden.de

Received 17 May 2009;
revised 3 August 2009;
accepted 14 October 2009

Abstract

Previous research suggests that patterns of comorbidity of common mental disorders among adults are best reflected by a hierarchical three-factor structure with two correlated factors ('anxious-misery' and 'fear') summarized in a second-order 'internalizing' factor and one 'externalizing' factor. This three-factor structure has not been examined yet in a sample of adolescents and young adults.

A representative sample of 3021 adolescents and young adults (baseline age 14–24) were prospectively followed over 10 years. Mental disorders were assessed according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) by using the standardized Munich Composite International Diagnostic Interview. Ten mental disorders (major depressive episode, dysthymia, generalized anxiety disorder, social phobia, specific phobia, agoraphobia, panic disorder, alcohol dependence, drug dependence, antisocial personality) were fitted to a series of Confirmatory Factor Analysis models using: (1) 12-month data, and (2) lifetime data from a person-year data set.

The three-factor model showed good fit to the observed data in our sample both when 12-month diagnoses and lifetime-to-date diagnoses from a person-year data file were used; yet the higher-order 'internalizing' factor summarizing 'anxious misery' and 'fear' had to be omitted.

The three-factor model could be replicated in a sample of adolescents and young adults with the exception that the second-order 'internalizing' factor was not consistent with the data. Further research is necessary to provide more complete insight into the structure of mental disorders by examining the stability of the structure of mental disorders in different developmental stages (ages) and by using a more extensive set of mental disorders.

Introduction

Comorbidity among mental disorders has been consistently shown to be the rule rather than the exception not just in clinical, but also in community samples among adults (Beesdo *et al.*, 2009a; Jacobi *et al.*, 2004; Kessler *et al.*, 2005; Wittchen and Jacobi, 2005) and youth (Beesdo *et al.*, in press; Fergusson *et al.*, 1993; Newman *et al.*, 1996; Wittchen *et al.*, 2000). In an attempt to account for this high comorbidity, Krueger (1999) fitted confirmatory factor analysis (CFA) models with one to four factors to 10 diagnoses (lifetime and 12-months) assessed in the cross-sectional National Comorbidity Survey (NCS) (Kessler *et al.*, 1994). Results suggested that patterns of comorbidity among common mental disorders in 15–54 year olds could be best accounted for in a hierarchical three-factor structure with two correlated factors 'anxious-misery' and 'fear' summarized in a second-order 'internalizing' factor and one 'externalizing' factor (Figure 1). 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, and alcohol dependence, drug dependence and antisocial personality disorder had high loadings on the 'externalizing' factor.

This proposed three-factor structure, comprising similar types and numbers of disorders, has been replicated in community studies of 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) (Table 1). Based on these findings, far-reaching implications have been discussed regarding basic research, pathogenesis, and diagnostic classification (Andrews *et al.*, 2009; Goldberg *et al.*, 2009; Krueger and Markon, 2006b; Watson, 2005; Watson *et al.*, 2008), suggesting that this model is not only phenotypically relevant, but may represent an important organizing devise for understanding common psychopathological processes.

However, despite the overall consistency in results when comorbidity patterns of adults are considered, there is also a range of observations and considerations that have stimulated considerable concerns about the appropriateness and utility of these findings. A number of methodological concerns have been raised regarding the factor analytic approach (Kessler *et al.*, 2005; Wittchen *et al.*, 1999ba) and the way the data were analyzed and interpreted (Wittchen *et al.*, 1999ba), particularly observations that the three-factor solution appears to be not robust against variations in diagnostic and sample

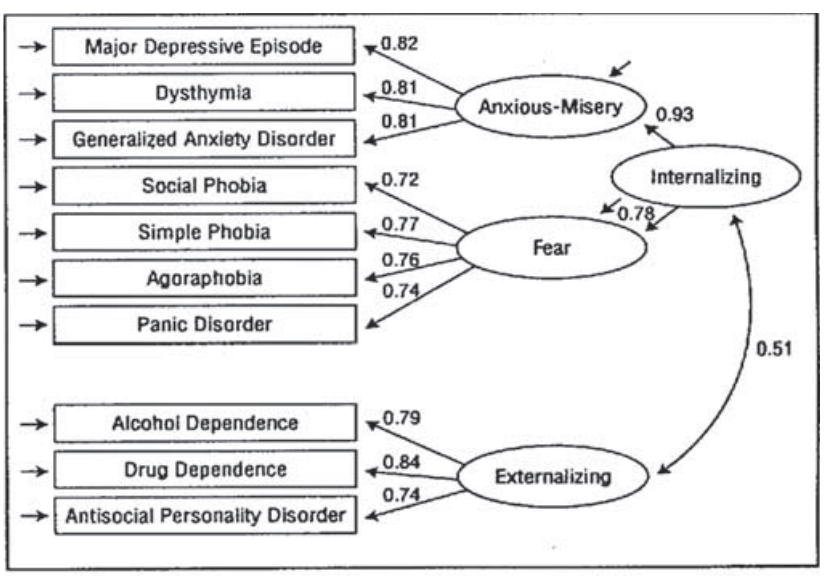

Figure 1 Krueger's three-factor model based on 10 lifetime disorders (NCS-data). (Figure from: Krueger, 1999, **56**, 921–926.)

Table 1 Overview of structural analyses performed in representative community sample datasets

Study (author)	Number of respondents	Age of respondents	Time	Criteria (instrument)	Special remarks	Number of disorders	Disorders/syndromes						
							MDE	Dys	GAD	Soc	Simp/Spec	Ago	Pan
NCS (Krueger, 1999)	8098	15–54	Lifetime/ 12-months	DSM-III-R (CIDI)	Cross-sectional	10	×	×	×	×	×	×	×
Replications													
NEMESIS (Vollebergh *et al.*, 2001)	7076 (5618 FU)	18–64 + one year FU	12-month	DSM-III-R (CIDI)	Longitudinal (one year interval), stability examined	9	×	×	×	×	×	×	×
NCS (Pt II subset) (Cox *et al.*, 2002)	5877	15–54	Lifetime/ 12-months	DSM-III-R (CIDI)	Cross-sectional, PTSD included	11	×	×	×	×	×	×	×
NCS (Pt II subset) (Watson, 2005)	5877	15–54	Lifetime	DSM-III-R (CIDI)	Cross-sectional, PTSD and bipolar included	12	×	×	×	×	×	×	×

Int. J. Methods Psychiatr. Res. 18(4): 204–220 (2009). DOI: 10.1002/mpr

OCD	PTSD	SAD	Alc Abu	Alc Dep	Drug Abu	Drug Dep	AAB/APD	Conduct	ODD	ADHD	IED	Neur	Man/Hypom	Analyses and models tested	Best fit/solution
				X		X	APD (lifetieme only) X							CFA One-factor Two-factor: Internalizing, Extenalizing Three-factor: Internalizing with anxious misery and fear, Externalizing; Four-factor: Affective, Anxiety, Substance use, Antisocial disorders	Three-factor-model – Internalizing with two subfactors: • anxious-misery: GAD, MDE, Dys • fear: Soc, Simp, Ago, Pan – Externalizing: Alc Dep, Drug Dep, APD Confirmed with lifetime and 12-month data, in random halfs of the sample and across men and women.
				X		X								CFA One-factor Two-factor: Internalizing, Extenalizing Three-factor: Internalizing with anxious misery and fear, Externalizing; Three-factor: Mood, Anxiety, Substance use disorders	Three-factor-model – Internalizing with two subfactors: • anxious-misery: GAD, MDD, Dys • fear: Soc, Simp, Ago, Pan – Externalizing: Alc Dep, Drug Dep Confirmed at both waves. (The structural stability of this model during a one-year period was substantial, and the differential stability of the three latent dimensions was considerable (higher though in externalizing than in internalizing, and higher in internalizing-fear than in internalizing-anxious-misery).)
	X			X		X	APD X							EFA	Three-factor model – Internalizing-fear: Pan, Ago, Soc, Simp – Internalizing-anxious-misery: MDE, Dys, GAD, PTSD – Externalizing: Alc Dep, Drug Dep, APD Confirmed for lifetime and 12-month diagnoses
	X			X		X	APD X						(Bip) X	Principal factor analysis	Three-factor-model – Internalizing-fear: Pan, Ago, Soc, Simp – Internalizing-anxious-misery: MDE, Dys, GAD, PTSD – Externalizing: Alc Dep, Drug Dep, APD – Bipolar Disorder fails to emerge as clear marker of any of these dimensions – it has weak and rel. identical loadings on all factors.

Table 1 *Continued*

Study (author)	Number of respondents	Age of respondents	Time	Criteria (instrument)	Special remarks	Number of disorders	Disorders/syndromes						
							MDE	Dys	GAD	Soc	Simp/Spec	Ago	Pan
ANSMHWB (Slade and Watson, 2006)	10641	>18	12-month	DSM-IV/ ICD-10	Cross-sectional	10 DSM/ 11 ICD	×	×	×	×		×	
DMHDS (Krueger *et al.*, 1998)	930 and 937	18 and 21	12-month	DSM-III-R (DIS)	Longitudinal (three year interval), stability examined	10	×	×	×	×	×	×	
NCS clinical sub-sample (Krueger and Finger, 2001)	251	15–54	Lifetime	DSM-III-R (CIDI)	Cross-sectional	7	×	×	×	×	×	×	×
NCS-R (Kessler *et al.*, 2005)	9282	>18	12-month	DSM-IV (CIDI)	Cross-sectional, many disorders considered	19 (17)	×	×	×	×	×	×	×

Int. J. Methods Psychiatr. Res. 18(4): 204–220 (2009). DOI: 10.1002/mpr

OCD	PTSD	SAD	Alc Abu	Alc Dep	Drug Abu	Drug Dep	AAB/APD	Conduct	ODD	ADHD	IED	Neur	Man/Hypom	Analyses and models tested	Best fit/solution
X	X			X		X						(ICD only) X		CFA One-factor Two-factor: Internalizing, Extenalizing Three-factor: Internalizing with subfactors distress and fear, Externalizing; Three-factor: Mood, Anxiety, Substance use disorders	Three-factor-model – Internalizing with two subfactors: • distress: MDD, Dys, PTSD, GAD, (Neur) • fear: Soc, Pan, Ago, OCD – Externalizing: Alc Dep, Drug Dep
X				X	(Mar) X	X	APD at age 21 X	at age 18 X						CFA One-factor Two-factor: Internalizing, Extenalizing Four-factor: Affective, Anxiety, Substance use, Antisocial disorders	Two-factor-model – Internalizing: GAD, Ago, Soc, Simp, OCD, MDE, Dys – Externalizing: Cond/APD, Mar Dep, Alc Dep To a significant extent, persons retained their relative positions on the latent factors across the three-year period from age 18 to 21.
														IRT	One-factor-model – Internalizing: all seven disorders (note: seven diagnoses measure the higher end of the factor; but lower half is not well reflected with these diagnoses)
X	X	X	X	X	X	X		X	X	X	X		X	EFA	*Note: EFA excluded disorders with negative correlations – OCD & SAD* Two-factor-model – Internalizing: MDE, Pan, Ago, Spec, Soc, GAD, PTSD – Externalizing: Conduct, Alc Abu, Alc Dep, Drug Abu, Drug Dep – five disorders had factor loadings of 0.30 or higher on both factors (Dys, mania/hypomania, ODD, ADHD, IED), but higher loadings on internalizing than externalizing factor
														LCA	LCA: seven factors Factor one: Unaffected respondents Factor two: Pure inernalizing Factor three: Pure externalizing Factor four: Comorbid internalizing Factor five: Comorbid internalizing and/or externalizing Factor six: Highly comorbid MDEs Factor seven: Highly comorbid Bipolars

Table 1 *Continued*

Study (author)	Number of respondents	Age of respondents	Time	Criteria (instrument)	Special remarks	Number of disorders	Disorders/syndromes						
							MDE	Dys	GAD	Soc	Simp/Spec	Ago	Pan
Meta-analysis (Krueger and Markon, 2006a)	23557	n.a. (15+)	n.a.	DSM-III-R/ DSM-IV	Includes twin study §	11	×	×	×	×	×	×	×

Disorders (Dx): AAB: Adult Antisocial Behavior; ADHD: Attention Deficit Hyperactivity Disorder; Ago: Agoraphobia; Alc Abu: Alcohol Abuse; Alc Dep: Alcohol Dependence; APD: Antisocial Personality Disorder; Bip: Bipolar Disorder; Cond: Conduct Disorder; Drug Abu: Drug Abuse; Drug Dep: Drug Dependence; Dys: Dysthymia; GAD: Generalized Anxiety Disorder; IED: Intermittent Explosive Disorder; Mar: Marijuana/Cannabis Dependence; MDE: Major Depressive Episode; Neur: Neurasthenia; OCD: Obsessive Compulsive Disorder; ODD: Oppositional Defiant Disorder; Pan: Panic disorder; PTSD: Posttraumatic Stress Disorder; SAD: Separation Anxiety Disorder; Simp/Spec: Simple or Specific Phobia; Soc: Social Phobia.
Studies: ANSMHWB: Australian National Survey of Mental Health and Well-Being; CHDS: Christchurch Health and Development Study; DMHDS: Dunedin Multidisciplinary Health and Development Study; NCS: National Comorbidity Survey; NCS-R: National Comorbidity Survey Replication; NEMESIS: Netherland Mental Health Survey and Incidence Study.
§ Meta-analysis includes: Virginia Twin Study (Kendler, 2003), NCS (Krueger, 1999), NCS-R (Kessler *et al*., 2005), NEMESIS (Vollebergh *et al*., 2001), DMHDS (Krueger *et al*., 1998).
Instruments: CIDI: Composite International Diagnostic Inverview; DIS: Diagnostic Interview Schedule.
Analysis: CFA: Confirmatory Factor Analysis; EFA: Exploratory Factor Analysis; IRT: Item Response Theory; LCA: Latent Class Analyses; MGA: Multivariate Genetic Analysis; SEM: Structural Equation Modeling.

composition: Other structures were found to fit the observed data when the number of included diagnoses was reduced (Krueger and Finger, 2001) or increased (Kessler *et al*., 2005), or when non-adult samples were used (Krueger *et al*., 1998) (compare Table 1). For example, in an unselected birth cohort of 18 and 21 year olds, a two-factor internalizing-externalizing model fit the data (Krueger *et al*., 1998). Number and type of disorders used in this study were identical to those in the original Krueger (1999) analysis except that obsessive compulsive disorder was included instead of panic disorder. Of note, in this 'early' structure study by Krueger *et al*. (1998), the fits of a one-factor, two-factor and four-factor model (comprising affective, anxiety, substance dependence, and antisocial disorders) were tested, but not a three-factor model.

The purpose of the current paper is to investigate whether the three-factor structure, which has up to now been examined only in adult samples, fits data from a prospective-longitudinal community study of adolescents and young adults. This is important because of the potentially far-reaching implications of structural findings for future diagnostic classifications and research. We use identical procedures including same number and types of diagnoses as Krueger (1999) for this replication. In addition, we will test the fit of a two-factor solution because the only other structure study in a non-adult sample (Krueger *et al*., 1998) suggested this solution, although without testing the three-factor model. Our analyses will use 12-month diagnoses as well as lifetime-to-date diagnoses by using a person-year data file.

Methods

Sample

The Early Developmental Stages of Psychopathology (EDSP) study (Lieb *et al*., 2000; Wittchen *et al*., 1998b; Wittchen *et al*., 1998c) assessed mental disorders in a representative sample of $N = 3021$ adolescents and young adults aged 14–24 years at baseline (T0). The study includes follow-up surveys (T1/T2/T3), a family history component (T0/T2/T3) and direct assessments of parents (T1/T3).

OCD	PTSD	SAD	Alc Abu	Alc Dep	Drug Abu	Drug Dep	AAB/APD	Conduct	ODD	ADHD	IED	Neur	Man/Hypom	Analyses and models tested	Best fit/solution
				X		X	AAB X	X						CFA One-factor Two-factor: Internalizing, Extenalizing Four-factor: Affective, Anxiety, Substance use, Antisocial disorders	Three-factor-model – Internalizing • Distress: MD, Dys, GAD • Fear: Ago, Soc, Spec, Pan – Externalizing: Alc Dx, Drug Dx, Cond, AAB

The baseline sample was drawn in 1994 from government registries (greater Munich area, Germany). The baseline (T0) 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 n = 2210 completed T3 (RR = 73.2%; interval since baseline: 7.3–10.6 years).

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).

Diagnostic assessment

Face-to-face interviews were conducted 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), 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 Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) diagnoses derived by the M-CIDI (Reed *et al.*, 1998; Wittchen, 1994; Wittchen *et al.*, 1998a).

The same 10 diagnoses as considered in the NCS-work by Krueger (1999) are considered here: Major Depressive Episode (MDE), Dysthymic Disorder, Generalized Anxiety Disorder (GAD), Panic Disorder with or without Agoraphobia, Agoraphobia with or without a history of Panic Disorder, Social Phobia, Specific Phobia, Alcohol Dependence, illicit Drug Dependence, and Antisocial Personality (ASP). With the exception of ASP, all diagnoses were obtained by using the M-CIDI/DSM-IV algorithms (Wittchen and Pfister, 1997) without applying hierarchical exclusion rules. For Social Phobia and Specific Phobia, the impairment criterion was applied only if respondents were aged ≥18 years at the respective assessment. Due to the age of the respondents in the EDSP (not all respondents were 18 years or older), ASP was defined as meeting criteria for Conduct Disorder (CD) or Antisocial Personality Disorder (APD). Using symptom lists reflecting DSM-IV criteria, respondent CD was assessed at T1 using parent reports (n = 1053) and at T2 from the respondents (n = 2548). APD was assessed at T2 from the respondents (n = 2548). Information on ASP is available on n = 2638 respondents with either T1 family assessment or T2 respondent assessment and missing for others. A discussion of our approach to analysis with missing data is described later.

Statistical analysis

Analysis focuses on the replication of Krueger's (1999) three-factor model in our sample of adolescents and young adults by means of CFA. We also fitted models with the three latent factors 'externalizing', 'anxious-misery' and 'fear' while omitting the second-order 'internalizing' factor since the original model yielded loadings of 'internalizing' on 'anxious-misery' close to one indicating that one can hardly separate these factors in our data (see later). As no prior study has tested a three-factor structure

in a non-adult sample, we also tested the fit of a two-factor internalizing-externalizing model following findings of the only other study examining youth (Krueger *et al.*, 1998). We did not test a one-factor model nor a four-factor model because they did not fit prior data both in youth and adult samples when using a similar range and type of diagnoses (compare Table 1).

Two sets of analyses were conducted using data from all four assessment points focusing on (1) 12-month data and (2) person-year data in which the measures of interest were lifetime-to-date diagnoses.

- 12-month data: The 12-month diagnoses from all four assessments were compiled in a dataset with 9007 observations [3021 (T0) + 1228 (T1) + 2548 (T2) + 2210 (T3) = 9007] reflecting the current disorder status for 14–34 year olds at the respective assessment. For ASP, only 3601 observations were available [1053 (T1 family assessment) + 2548 (T2)].
- Person-year data: A person-year data file was build 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). For instance, a proband aged 20 at the last completed assessment contributes to the data set 20 observations. In cases where reported age-of-onset for a particular disorders varied for a given respondent over time, the minimum reported age-of-onset was used. The dataset had a total of 75,613 observations. Age-of-onset of ASP was defined as the parental onset information provided at T1 family assessment for the CD cases at that assessment, and for all other cases it was defined as 13 (APD requires pre-existing CD symptoms). Thus, 8488 observations with missing values occurred due to ASP only being assessed at T1 family assessment and T2 and a further 865 missing values occurred due to lacking age-of-onset information for specific diagnoses, leaving 66 260 complete observations.

Consistent with Krueger (1999) the analysed matrix in CFA was the tetrachoric correlation matrix of the disorders and the standardized root mean square residual RMS (usually abbreviated as SRMR; we use the abbreviation 'RMS' as in 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 to model misspecification and less sensitive than other global fit measures to distribution and sample size in covariance structure models (Hu and Bentler, 1998). We also calculated the CFI (Comparative Fit Index) and TLI (Tucker–Lewis Index; Relative Noncentrality Index) measures of model fit to confirm that the selection of the best-fitting model based on RMS was not sensitive to the particular measure of fit used. Values close to one in the CFI and TLI indicate good fit and values around 0.9 indicate acceptable fit. TLI might become greater than one. Finally, the chi-squared goodness of fit test is reported (irrespective of the limited value in very large samples).

CFA was carried out in Mplus, Version 5. Krueger (1999) used PRELIS and LISREL and seemingly has excluded cases with missing data. For each of the datasets we ran three analyses dealing differently with missing values: (a) all available information was used with the full information maximum likelihood method based on the missing at random assumption, (b) complete case analysis, and (c) missing values were replaced with zero; here, missing values due to missing age-of-onset information were not coded as zero because they are known to be cases. Since Mplus does not calculate RMS in case (a) (because the therefore necessary residuals are missing for some cases) and we wanted to ensure that our results were robust against choosing (a), (b) or (c) all three analyses were run.

Unlike in the Krueger (1999) analysis, our data were clustered within persons (because of our longitudinal data with four assessments). Simple weighted least squares yielded non-positive definite covariance matrixes between the latent variables here and the method was therefore not used. Instead, the weighted least squares mean and variance adjusted estimator based on a diagonal weight matrix was applied (known to yield more robust results for clustered data). Valid chi-square tests of model fit and standard errors taking into account clustering within persons could not be calculated in (b) and (c) because RMS can not be computed at the same time in Mplus.

At T0 the probands had been sampled with different weights according to age (Wittchen *et al.*, 1998b; Wittchen *et al.*, 1998c). In all analyses sampling weights were used to adjust the sample at T0 to the source population regarding age, sex and geography.

Results

Replication using 12-month data

Table 2 reveals the results of the CFA using 12-month diagnoses. In our primary analyses we tested Krueger's three-factor model with the second-order internalizing factor (column one). Both when coding missing values of ASP as zero and in the complete case analysis, fit was good

Int. J. Methods Psychiatr. Res. 18(4): 204–220 (2009). DOI: 10.1002/mpr

Table 2 Results of CFA using 12-month diagnoses

	Three factors with 'internalizing'	Three factors without 'internalizing'	Two factors 'internalizing' and 'externalizing' only
Missings of ASP treated with the full information maximum likelihood method			
N	9007	9007	9007
Model fit			
Number of free parameters	23	23	22
Chi-square	24.44	24.44	33.63
Degrees of freedom	21	21	22
p-Value	0.272	0.272	0.054
Fit indices			
CFI	0.996	0.996	0.988
TLI	0.997	0.997	0.989
RMS	*n.a.*	*n.a.*	*n.a.*
Subsample with complete data			
N	3601	3601	3601
Model fit			
Number of free parameters	13	13	11
Chi-square	12.82[1]	12.82[1]	15.91[1]
Degrees of freedom	19	19	19
p-Value	0.848	0.8478	0.6631
Fit indices			
CFI	1.000	1.000	1.000
TLI	1.014	1.014	1.007
RMS	0.060	0.061	0.066
Missings of ASP coded zero			
N	9007	9007	9007
Model fit			
Number of free parameters	13	13	11
Chi-square	31.55[1]	31.55[1]	43.76[1]
Degrees of freedom	24	24	25
p-Value	0.139	0.1387	0.0115
Fit indices			
CFI	0.995	0.995	0.987
TLI	0.994	0.994	0.986
RMS	0.044	0.044	0.050

[1] Tests not valid because it does not take clustering within probands into account.

with RMS equalling 0.044 and 0.060, respectively; also CFI and TLI values indicated good fit. Fit indices and also factor loadings (table available on request) showed very similar results across all three models.

Figure 2(a) depicts results for the model with missing values coded zero. Of note, all models yielded high factor loadings of 'anxious-misery' on 'internalizing' (>0.9). Considering the standard error this value could be greater than 1.0. Therefore we also tested the three-factor model without 'internalizing' (Table 2, middle column). Model fits and factor loadings were almost identical to the ones of the three-factor model with 'internalizing'. Figure 2(b) depicts results for the model with missing values coded zero. High factor loadings were found for 'anxious misery' with 'fear' (0.834–0.848); estimates for 'externalizing' with 'anxious misery' (0.456 – 0.505) and 'fear' (0.453–0.494) were lower. In both three-factor models (with and without 'internalizing'), factor

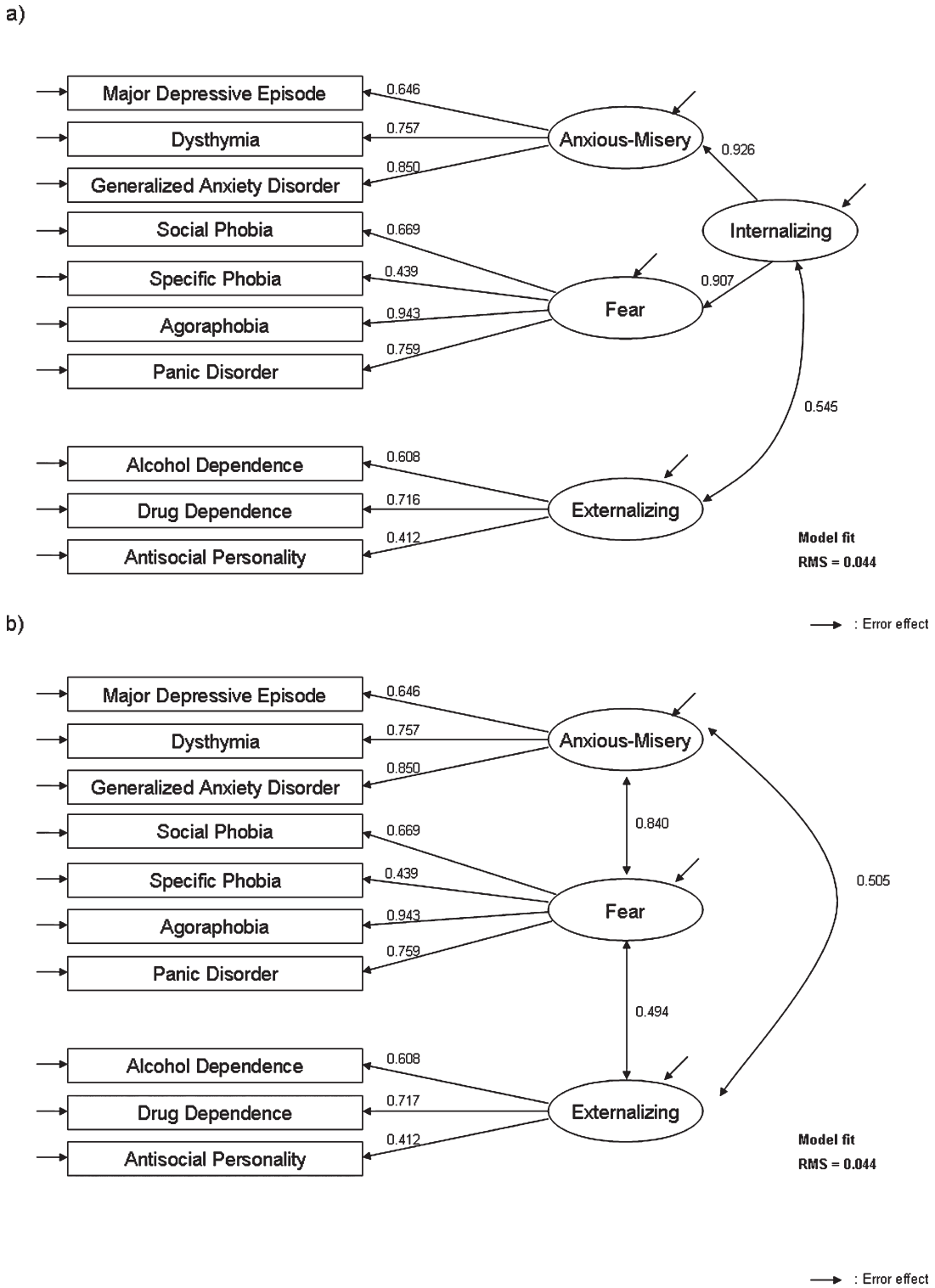

Figure 2 Results of CFA using 12-month diagnoses ($N = 9007$ cases, missing values of antisocial personality coded as zero): (a) three-factor model with 'internalizing', (b) three-factor model without 'internalizing'.

Int. J. Methods Psychiatr. Res. 18(4): 204–220 (2009). DOI: 10.1002/mpr

loadings for the disorders were generally high (>0.6) except of the factor loading of 'fear' on specific phobia (0.427 to 0.439).

The two-factor internalizing-externalizing model also fit the data satisfactorily but worse than the three-factor models (Table 2, third column).

Replication using person-year data

Table 3 shows the results of the CFAs using person-year data and focusing on lifetime-to-date disorders. The three-factor Krueger model provided a good fit indicated by RMS (0.057 and 0.061), CFI and TLI values (see Table 3). Again, 'internalizing' loaded almost perfectly on

Table 3 Results of CFA using person-year data

	Three factors with 'internalizing'	Three factors without 'internalizing'	Two factors 'internalizing' and 'externalizing' only
Missings of ASP treated with the full information maximum likelihood method			
N	75 613	75 613	75 613
Model fit			
Number of free parameters	23	23	21
Chi-square	49.287	49.250[1]	60.089[1]
Degrees of freedom	22	22	23
p-Value	<0.001	<0.001	<0.001
Fit indices			
CFI	0.984	0.985	0.979
TLI	0.984	0.984	0.979
RMS	*n.a.*	*n.a.*	*n.a.*
Subsample with complete data			
N	66 260	66 260	66 260
Model fit			
Number of free parameters	13	13	25
Chi-square	–[1]	621.52[2]	738.93[2]
Degrees of freedom	24	24	25
p-Value	–[1]	<0.001	<0.001
Fit indices			
CFI	–[1]	0.972	0.966
TLI	–[1]	0.970	0.966
RMS	–[1]	0.057	0.061
Missings of ASP coded as zero			
N	74 634	74 634	74 634
Model fit			
Number of free parameters	13	13	11
Chi-Square	726.94[2]	726.95[1]	846.86[1]
Degrees of freedom	25	25	26
p-value	<0.001	<0.001	<0.001
Fit indices			
CFI	0.971	0.971	0.966
TLI	0.970	0.970	0.966
RMS	0.058	0.058	0.061

[1] The covariance matrix of the latent variables was not positive definite.
[2] Tests not valid because they do not take clustering within probands into account.

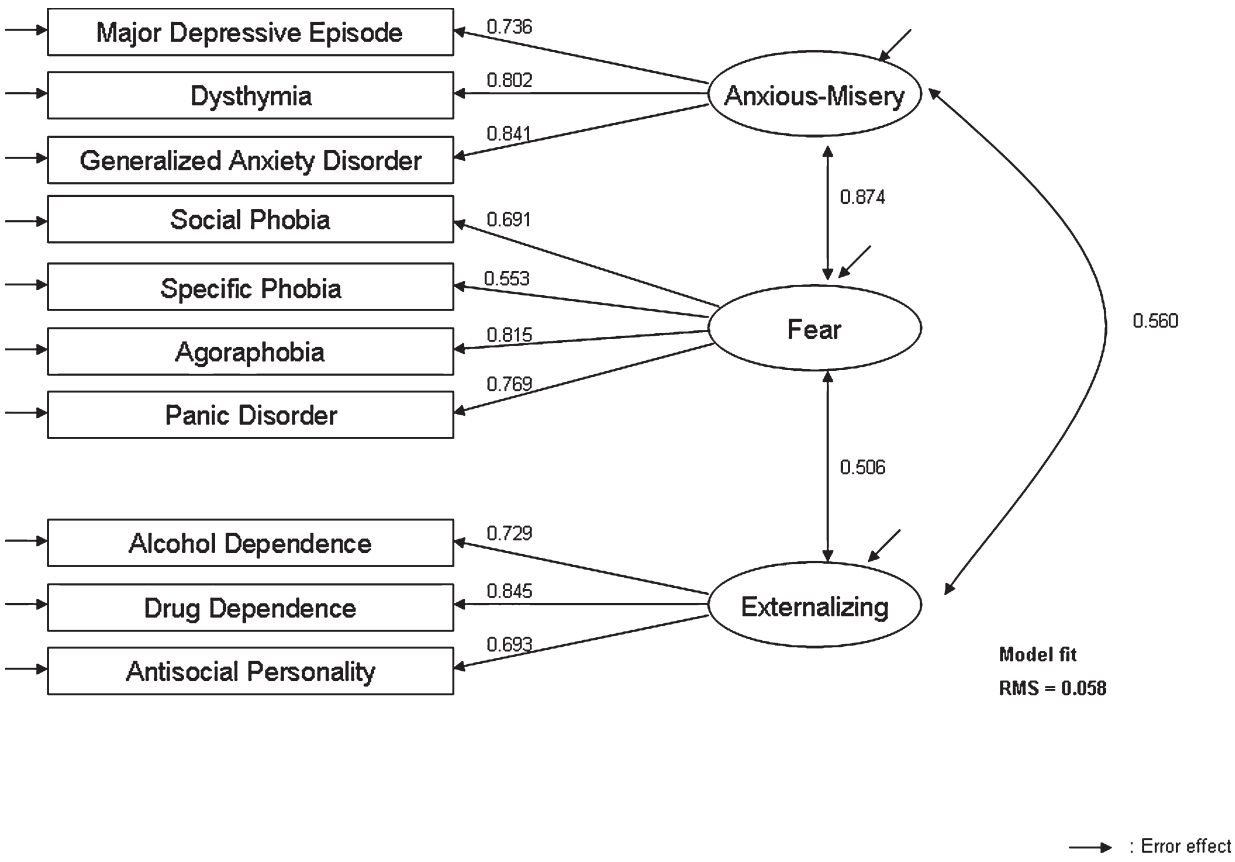

Figure 3 Results of CFA using person-year data (N = 74 634 cases, missing values of antisocial personality coded as zero).

'anxious-misery' (0.972 when imputing missing values and 0.982 when coding them zero). Considering the standard errors these values could be greater than one. In the complete data analysis, the second-order factor 'internalizing' was even estimated to be greater than one yielding a negatively definite covariance matrix between the latent variables meaning that the specified model is logically inconsistent with the covariance structure of our data.

Omitting 'internalizing' from the three-factor model (Table 3, middle column) again revealed almost identical fit indices compared to the models with 'internalizing'. Model-fit was also good in all models. Consistently, high factor loadings were found for 'anxious-misery' with 'fear' (0.861–0.881); estimates for 'anxious-misery' with 'externalizing' (0.557–0.580) and 'fear' with 'externalizing' (0.488–0.501) were lower (results available on request). Figure 3 depicts results for the model with missing values coded zero.

We finally also tested the two-factor internalizing-externalizing model to our person-year data. Again, model fit was satisfactory but worse than for the three-factor models (Table 3, third column).

Discussion

The consistent observation of high comorbidity among mental disorders even in community samples (e.g. Kessler *et al.*, 2005; Wittchen and Jacobi, 2005) has prompted investigations of higher order structure. Using only 10 diagnoses assessed among adults in the cross-sectional NCS, Krueger (1999) suggested that the bivariate associations (tetrachoric correlations) between all logically possible pairs of lifetime and 12-month disorders assessed in his study could be adequately characterized as caused by a hierarchical three-factor structure with 'anxious-misery' and 'fear' as correlated factors summarized in a second-order 'internalizing' factor and one

'externalizing' factor. This three-factor model has been replicated in several other community adult samples using both lifetime (Cox *et al.*, 2002; Watson, 2005) and 12-month diagnoses (Cox *et al.*, 2002; Slade and Watson, 2006; Vollebergh *et al.*, 2001). To date, though, only one other study examined the factor structure specifically among adolescents and young adults (Krueger *et al.*, 1998). In this study, a two-factor internalizing-externalizing model fit the data better than a one-factor and four-factor model; a three-factor model was not tested.

The present study is the first that tested the original Krueger (1999) three-factor model in a sample of adolescents and young adults from a prospective-longitudinal community survey. We also examined the two-factor solution that was found to have good model fit in the only other study that examined a non-adult sample (Krueger *et al.*, 1998).

We used similar methods and conventions as in the Krueger (1999) analysis of the NCS data. Fitting identical 10 diagnoses from 14–34 year olds to CFAs, the current study provides replication for the three-factor model using both 12-month data and person-year data, yet with the difference that the higher order 'internalizing' factor summarizing 'anxious-misery' and 'fear' had to be omitted indicating that there is no hierarchy between the three factors 'anxious misery', 'fear' and 'externalizing' in our data. A two-factor internalizing-externalizing model fitted worse to the data than the three-factor model.

Our findings among adolescents and young adults are consistent with data from adults (Cox *et al.*, 2002; Slade and Watson, 2006; Vollebergh *et al.*, 2001; Watson, 2005), in that the Krueger three-factor model could be replicated. In contrast to other studies, however, the second-order 'internalizing' factor was not consistent with our data. In Krueger (1999) factor loading of 'anxious-misery' on 'internalizing' was also high (0.93), while the loading of 'fear' on 'internalizing' was somewhat lower (0.78). Moderate changes in design might have yielded data with a 'more than perfect' loading that did not allow to fit the three factor model including 'internalizing' in our data. Our prospective-longitudinal study may have yielded higher comorbidities than data sets in the Krueger model fitted before (cross-sectional 12-month diagnoses (Cox *et al.*, 2002; Slade and Watson, 2006; Vollebergh *et al.*, 2001) and lifetime diagnoses derived from a single assessment (Cox *et al.*, 2002; Krueger, 1999; Watson, 2005).

Noteworthy, some of the factor loadings in our data were also much smaller than in the Krueger model (1999) where they range from 0.72 to 0.84. This is particularly evident for 'fear' on specific phobia (models with 12-month data: 0.439, person-year data: 0.553 versus 0.77 in Krueger's model). This indicates that 'fear' is less strongly related to specific phobia in our sample and thus specific phobia seems to be a less reliable indicator for this factor. In general we found more heterogeneity between loadings of latent variables on diagnoses.

Beyond the three-factor model, we also tested a two-factor 'internalizing'-'externalizing' model following findings from the only other non-adult study (Krueger *et al.*, 1998). Our findings are consistent with this study in that the two-factor model fitted the data. In contrast to this other study, we also tested a three-factor model and showed that the three-factor solution fitted the data better. It remains open, whether these findings would also have occurred in this other non-adult sample if tested.

Before implications of these findings can be discussed several limitations must be considered. We replicated the Krueger (1999) findings using the same methods and assumptions. Yet, if these methods and assumptions are inadequate, the model is questionable. Thus, a first group of limitations relates to the question to what degree such data are appropriate to identify underlying dimensions and core psychopathological processes.

(1) Identical to Krueger (1999) and others (Cox *et al.*, 2002; Krueger *et al.*, 1998; Krueger and Markon, 2006a; Slade and Watson, 2006; Vollebergh *et al.*, 2001; Watson, 2005) we only examined a quite restricted small set of disorders not including other anxiety diagnoses (specific phobia subtypes, obsessive compulsive disorder, post-traumatic stress disorder, separation anxiety), other affective disorders (bipolar disorders, respectively hypomania and mania), substance abuse disorders and a range of other clinical conditions (eating, somatoform, psychotic disorders). Thus, the model only applies to the few diagnoses examined and not to the entire set of DSM-IV conditions. There is indication that this structure does not hold or that alternative models are preferable when the number of included diagnoses is increased (Kessler *et al.*, 2005; Wittchen *et al.*, 2009).
(2) The data and the analyses only reflect associations for threshold disorders with the assumption that they cluster together temporally, which is questionable. For example, even for the 12-month analyses the co-occurrence of anxiety and depression might mean in one case one disorder is always present within a 12 month window, but the other only for one month. In other cases, there might be complete temporal overlap, or one disorder is only present on a subthreshold level while the other one meets the threshold level. This is

due to the inherent inaccuracy of the assessment instrument that cannot ascertain fine graded temporal patterns for most disorders. Further, one needs to take into account recall bias, particularly in regard to age-of-onset. This is, however, less concerning for the present prospective multi-wave study of a young age cohort in contrast to adult cross-sectional studies. Nevertheless, this type of data and analyses might not reflect the clinically relevant more subtle distinctions of psychopathological variability and their impact. The failure to distinguish between concurrent and lifetime comorbidity might result in flawed conclusions (Knäuper and Wittchen, 1994; Wittchen, 1996).

(3) The examination of the higher order internalizing factor in previous and the present analyses is not developmentally sensitive. There is substantial evidence that the meaning (Knäuper and Wittchen, 1994), the incidence/age-of-onset patterns (Beesdo *et al.*, in press, 2007; Wittchen *et al.*, 1999b), the persistence and expression of mental disorders and their overlap is quite different in children, adolescents, young and older adults (Beesdo *et al.*, 2009b; Kessler *et al.*, 2005). We did not examine whether the three-factor structure is stable across development. Notwithstanding our replication of Krueger's (1999) model in our sample of adolescents and young adults, other solutions may be appropriate in other samples or age groups. Given the differences in the age-of-onset patterns of mental disorders with anxiety disorders most frequently occurring in childhood and depressive disorders and substance use disorder showing a later onset in adolescence or adulthood, it is conceivable that different factor solutions are reflective of the structure of common mental disorders in different developmental stages. Samples with broad age compositions might fail to reflect the potentially considerable differences.

A second group of limitations refers to the appropriateness of the statistical procedures and the degree to which they allow for far reaching interpretations.

(1) In his original work Krueger (1999) only examined four models when deciding on the candidate dimensions. 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).

(2) The use of tetrachoric correlations implicitly assumes that the joint distribution underlying two disorders is bivariate normal – the latent variables being linearly und uniformly related to one another (Kraemer, 1997). This assumption is likely to be violated, given the type of data used. Kraemer (1997, p. 1121) 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.

To conclude, we replicated – yet while discarding the second-order 'internalizing' factor – the Krueger (1999) three-factor model in a sample of adolescents and young adults. However, the consistency to which this model can be replicated is only one essential consideration. Particularly in light of the substantial degree of recent speculation and implications attached to this model (Andrews *et al.*, 2009; Goldberg *et al.*, 2009), further research is necessary to provide more complete insight into the meaning of these findings and their implications particularly for diagnostic and classificatory issues (Wittchen *et al.*, 2009). A more comprehensive examination of developmental issues and stabilities, as well as a broader coverage of disorders seems to be urgently needed to assess the robustness of this model (see Wittchen *et al.*, in this issue).

Acknowledgements

This work is part of the Early Developmental Stages of Psychopathology (EDSP) Study and is funded by the German Federal Ministry of Education and Research (BMBF) project no. 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, 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.

Beesdo 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 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 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*.

Beesdo K., Bittner A., Pine D.S. *et al.* (2007) Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Archives of General Psychiatry*, **64**(8), 903–912.

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.

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.

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.

Kendler K.S., Prescott C.A., Myers J., Neale M.C. (2003) The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, **60**, 929–937.

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.

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.

Knäuper B., Wittchen H.-U. (1994) Diagnosing major depression in the elderly: Evidence for response bias in standardized diagnostic interviews? *Journal of Psychiatric Research*, **28**, 147–164.

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., Finger M.S. (2001) Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assessment*, **13**, 140–151.

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.

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.

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.

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. (1996) Critical issues in the evaluation of comorbidity of psychiatric disorders. *British Journal of Psychiatry*, **168**, 9–16.

Wittchen H.-U., Beesdo K., Gloster A.T. (2009) The position of anxiety disorders in structural models of mental disorders. *Psychiatric Clinics of North America*, **32**, 465–481.

Wittchen H.-U., Beesdo K., Gloster A.T. (in press) A new metastructure of mental disorders: a helpful step into the future or a harmful step back to the past? *Psychological Medicine*, **39**, 2083–2089.

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 J.L.), 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*, **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.