

Michael Erhart
Manfred Döpfner
Ulrike Ravens-Sieberer
and the BELLA study group

Psychometric properties of two ADHD questionnaires: comparing the Conners' scale and the FBB-HKS in the general population of German children and adolescents – results of the BELLA study

Members of the BELLA study group: Ulrike Ravens-Sieberer (Principal Investigator), Claus Barkmann, Susanne Bettge, Monika Bullinger, Manfred Döpfner, Michael Erhart, Beate Herpertz-Dahlmann, Heike Hölling, Franz Resch, Aribert Rothenberger, Michael Schulte-Markwort, Nora Wille, Hans-Ulrich Wittchen.

Dr. M. Erhart
Prof. Dr. U. Ravens-Sieberer, MPH (✉)
University Clinic Hamburg-Eppendorf
Centre for Obstetrics and Paediatrics
Dept. of Psychosomatics in Children
and Adolescents
Martinistraße 52, W29
20246 Hamburg, Germany
Tel.: +49-40/42803-7378 or -7585
Fax: +49-40/42803-5105
E-Mail: ravens-sieberer@uke.uni-
hamburg.de

Prof. Dr. M. Döpfner
Dep. Child and Adolescent Psychiatry
University of Cologne
Cologne, Germany

■ **Abstract** *Background and objective* To examine and compare the psychometric properties of two short screening instruments for children and adolescents suffering from attention deficit-/ hyperactivity disorder (ADHD). The Conners' Hyperactivity Index consists of ten items that assess symptoms of hyperactivity through self-report and parents' proxy. The German ADHD Rating scale (FBB-HKS/ADHS) consists of 20 items that assess the severity and perceived burden of inattention, hyperactivity, and impulsiveness as defined by the ICD-10 and DSM-IV. *Methods* Within the BELLA module of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), the parents of 2,863 children and adolescents rated the Conners' Hyperactivity Index and the FBB-HKS. *Results* The internal consistency of item responses was assessed via Cronbach's α and showed that both instrument scores were able to obtain a reliable measurement. The factorial validity of the FBB-HKS measurement model as well as the unidimensionality of the Conners' scale was tested by means of exploratory and confirmatory factor analysis (EFA and

CFA), indicating satisfactory goodness of fit for the FBB-HKS (RMSEA = 0.06) and some deviation from the unidimensionality assumption of the Conners' scale. Stability of results across age could be confirmed with few exceptions. Mean scores differences were found between both sexes, age groups, and different socioeconomic status groups (Winkler-Index) with males, younger respondents, and children with low socioeconomic status displaying more ADHD-related behaviour. Correlation coefficients between the two instruments' scores and other scales assessing emotional and behavioural problems hinted at convergent validity. *Conclusion* Both instruments' scores showed reliability as well as factorial and convergent / discriminant validity. The pros and cons of the two instruments as well as for which purpose and under which circumstances one of the measures can be favoured must be considered prior to applying such a measure.

■ **Key words** attention deficit-/ hyperactivity disorder – parent-report – screening instruments – reliability – factorial validity

Introduction

An estimated proportion of 5–10% of school-aged children worldwide are affected by attention deficit-/hyperactivity disorder (ADHD, DSM-IV), also known as hyperkinetic disorder (HD, ICD-10) [4]. This prevalent disorder affects the child's sense of well-being, self-worth, resilience, overall health-perception, and psychosocial functioning [13, 28, 38, 39]. Furthermore, the family and community are burdened. Together, these aspects might create one of the greatest aggregate burdens of suffering of all mental disorders that have an onset in childhood [5].

Therefore, the assessment of ADHD becomes a major issue in clinical and epidemiological studies as well as in routine clinical services. The currently available set of ADHD rating scales encompasses several instruments to assess this disorder as defined by the DSM-IV or the ICD-10. These instruments are based on information from parents, teachers, or the self-report of the children or adolescents [21]. The test-theoretical quality of many of these instruments was confirmed in studies that were mainly conducted in English-speaking countries [14].

Several ADHD symptom checklists based on DSM-IV criteria were developed and analysed in English-speaking countries [e.g., 24, 40]. The German ADHD Rating scale (FBB-HKS/ADHS) includes 20 items describing the symptom criteria of both the ICD-10 and DSM-IV as well as additional items assessing symptom onset, symptom duration, pervasiveness, and functional impairment. The items can be rated by parents or teachers. The FBB-HKS/ADHS is part of the comprehensive diagnostic system for mental disorders in childhood and adolescence (DISYPS-KJ, [22]; DISYPS-II, [23]). Additionally, Breuer and Döpfner [8, 9] developed and analysed an adaptation of the FBB-HKS for preschool children. Exploratory factor analyses of parent-reports identified two components measuring the dimensions of inattention and hyperactivity-impulsivity, which is in line with the DSM-IV [10, 23, 26]. For preschool children, these 2-dimensions could be replicated in confirmatory factor analyses [9]. The validity of this measurement model encompassing two factors could be shown in confirmatory factor analysis with other ADHD symptom rating scales based on DSM-IV samples from English-speaking countries (Wolraich et al. [43]), and also in a pan-European sample of children and adolescents with ADHD symptoms [21].

The 18-symptom criteria for attention deficit-/hyperactivity disorders according to the DSM-IV are similar to the 18 symptom criteria for hyperkinetic disorder according to the ICD-10. However, in the ICD-10, hyperactivity and impulsivity are concep-

tualised separately by five and four criteria, respectively. In the cross-national examination of Döpfner et al. [21], three-factor solutions were extracted to test whether the hyperactivity-impulsivity factor can be broken down to a separate hyperactivity factor and impulsivity factor. Such a three-factor solution with separate inattention, hyperactivity, and impulsivity factors was observed in four of seven national subsamples and for boys but not for girls. Moreover, empirical evidence for the validity of three-factor solutions comprising inattention, hyperactivity, and impulsivity separately could be found in studies with the FBB-HKS in German samples [27].

Previous studies [27] showed that the items of the subscales inattention, hyperactivity, and impulsivity as well as hyperactivity-impulsivity and the total score were answered in an internally consistent manner. Cronbach's α was satisfactory to very good, ranging from 0.78 to 0.90. In their pan-European study using another ADHD rating scale, Döpfner et al. [21] reported satisfying results on internal consistency for the two scales inattention ($\alpha = 0.81$) and hyperactivity-impulsivity ($\alpha = 0.87$).

The Conners' Hyperactivity Index [16] assesses the occurrence of the most important symptoms of ADHD within the last month. Ten items are provided and answered by the children or adolescents themselves or their parents. The initial Conners' Parent Rating Scale (CPRS) was developed as a comprehensive checklist for acquiring parental reports on the basic problems present in children who had been referred to an outpatient psychiatric setting [15]. In its original form, the CPRS contained items grouped in terms of sleep problems, eating problems, temper problems, problems with keeping friends, school-related problems, etc. Later, an "additional" problems category was added that included items covering the cardinal symptoms of ADHD: hyperactivity, impulsivity, and inattention. Some factor analytic research with the CPRS on clinical samples suggested slightly different CPRS factor structures than were originally reported. Despite these differences, good reliability of the CPRS, as assessed by test-retest and internal consistency reliability [17], could be shown. In addition, the CPRS's concurrent validity is well established by high correlations with similar factors on other parent rating scales, such as the Child Behavior Checklist [1, 34] and the behavior problem checklist [2, 12]. The Conners' scale has been widely used in clinical and epidemiological studies. Previous studies showed the Conners' scale to be able to distinguish between subjects differing in hyperactivity but not between those differing in inattention [41]. Other studies have shown a weak discrimination between subtypes of ADHD [7].

This paper aims to compare the psychometric properties of the Conners' Hyperactivity Index and the FBB-HKS. Psychometric properties under study are the factorial validity of both instruments' measurement model, the internal consistency of item answers, and the reliability of the test scores.

Methods

■ Design and sample

Conceptualisation, design, and procedure of the Mental Health Module (BELLA study) are described in detail in Ravens-Sieberer et al. [37]. The participants of the BELLA study were randomly recruited from the national, representative sample of 17,641 families participating in the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) conducted by the Robert Koch-Institute. The KiGGS and the BELLA survey took place between May 2003 and May 2006 in 167 cities and communities representative of Germany. The overall response rate was 66.6% (KiGGS). A random selection of 4,199 families from the KiGGS sample with children aged 7–17 years were asked to participate in the BELLA study. Of these eligible families, 70% agreed to participate and 68% (1,389 girls and 1,474 boys) could be surveyed. From the 2,863 families participating in the BELLA study, 1,142 had children aged 7–10 years, 780 had children aged 11–13 years, and 941 had children aged 14–17 years. In each family, one parent was questioned with a standardised computer assisted telephone interview (CATI). Children aged 11 years and older were questioned as well. In addition, the participants were asked to fill in a mailed paper and pencil questionnaire. Sample data were weighted to correct for deviation of the sample from the age-, gender-, regional-, and citizenship-structure of the German population (reference data 31 December 2004).

■ Instruments

The German ADHD Rating scale (FBB-HKS/ADHS, [22]) includes 20 items of the symptom criteria of both the ICD-10 and DSM-IV as well as additional items assessing symptom onset, symptom duration, pervasiveness, and functional impairment. Parents indicated the frequency of each item statement or symptom on a 4-point answer scale ranging from never or rarely (0) to very often (3), with higher scores indicating greater ADHD-related behaviour. The mean item score is calculated for every dimension. It is possible to classify respondents with regards

to the criteria for ADHD subtypes as defined by the DSM-IV (predominantly inattentive, predominantly hyperactive-impulsive, combined, e.g., six of nine symptoms of inattention rated as "often" or "very often" and/or six of nine symptoms of hyperactivity-impulsivity rated as "often" or "very often").

The Conners' ten-item Hyperactivity Index [16] assesses the occurrence of the most important symptoms of ADHD within the last month. The ten items were rated by the parents on a 4-point answer scale ranging from 0 for not at all true to 3 for very much true. The raw sum scores were transformed into T-values with a mean of 50 and a standard deviation of 10.

The strengths and difficulties questionnaire [25] was applied as a brief behavioural screening questionnaire that assesses positive or negative attributes by 25 items focusing on the following dimensions: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behaviour. Each of the 25 items of the SDQ is scored on a 3-point scale (0 = not true, 1 = somewhat true, or 2 = certainly true), with higher scores indicating larger problems. The prosocial behaviour scale was not used for the current analysis. Items of the four problem areas are summed up to generate a total difficulties score (0–40). The sub-scores of the four difficulties dimensions range from 0 to 10.

Socioeconomic status (SES) was assessed with the Winkler-Index [42], which classifies the families of the respondents into those with low, medium, and high SES, taking income, education, and parental working position into account.

■ Statistical analyses

The statistical analyses are based on the weighted sample data to represent the age-, gender-, regional-, and citizenship-structure of the German population (reference data 31 December 2004). The number of cases reported in tables and in text refers to weighted data and, thus, might deviate from the number of cases reported in the former description of the sample.

The factorial validity of the FBB-HKS measurement model was tested by means of exploratory principal component analysis and confirmatory factor analysis. Using the LISREL 8 software, a confirmatory factor analysis was conducted by specifying a linear structural equation model [31] according to the multidimensional measurement model of the FBB-HKS. Similarly, a unidimensional factor model was specified for the Conners' scale to test the unidimensionality assumption of the instrument. Identifiability of the model parameters was ensured by each observed variable loading onto only one latent construct and by

fixing the variance of each latent variable to one. The succeeding complete standardisation of the model enabled correct parameter estimates [35]. The database for the unweighted least squares (UWLS) estimation of the model parameters was the polychoric correlation matrix of the observed indicators. As the UWLS estimation procedure does not require multivariate normal distribution of the data, no a priori normalisation of the observed variables was applied. For cases with less than 20% missing values on the SDQ items, missing values were replaced by the multiple imputation expected maximisation (EM) procedure of PRELIS 2 [31]. The goodness of fit of the model was assessed by the root mean square residual (RMSEA) and the adjusted goodness of fit index (AGFI). A RMSEA less than 0.06 (0.08) was taken as an indicator of excellent (adequate) fit between the specified model and the data [29].

The internal consistency of item responses was assessed via Cronbach's α for every measurement domain. Mean scores were calculated. Mean score differences between males and females, age groups (11–13 vs. 14–17 years), and low, medium, and high

socioeconomic status (Winkler-Index) were examined with ANOVA. Correlation coefficients were calculated between the two instruments and other scales assessing emotional and behavioural problems in order to assess convergent and discriminant validity.

Results

Factorial validity—exploratory and confirmatory factor analysis

A principal component analysis of the FBB-HKS inter-item correlation matrix was conducted. The first four unrotated principal components had eigenvalues greater than one (7.27, 1.92, 1.05, 1.04). In a first step, two principal components were extracted and rotated to simple structure according to the direct oblimin criteria. Then, a three-component solution was examined. The two-component structure resembled the theoretical measurement model with the dimensions of inattention and hyperactivity impulsivity. A total of 46% of the variance could be explained. Pat-

Table 1 Pattern coefficients: exploratory principal component analysis (PCA) with direct oblimin rotation and confirmatory factor analysis (CFA) using structural equation modeling of the FBB-HKS inter-item correlation matrix

Dimensions and items (abbreviated)	Two-component solution				Three-component solution					
	1st		2nd		1st		2nd		3rd	
	PCA	CFA	PCA	CFA	PCA	CFA	PCA	CFA	PCA	CFA
Inattention										
No close attention to details	-0.06	0.31	0.78	0.68	0.81	0.68	-0.15	0.32	0.09	0.25
Difficulty sustaining attention	0.18	0.38	0.64	0.75	0.64	0.75	-0.12	0.39	0.09	0.31
Does not seem to listen	0.21	0.27	0.48	0.60	0.50	0.60	0.04	0.28	0.20	0.22
Fails to finish work	0.02	0.32	0.76	0.71	0.74	0.71	0.10	0.33	-0.08	0.26
Has difficulty organising tasks and activities	-0.01	0.27	0.68	0.58	0.67	0.58	0.05	0.28	-0.06	0.22
Avoids tasks that require mental effort	0.07	0.40	0.73	0.72	0.73	0.72	0.06	0.41	0.02	0.33
Loses things necessary for tasks or activities	-0.05	0.25	0.68	0.55	0.67	0.55	0.02	0.26	-0.08	0.21
Is easily distracted	0.18	0.38	0.65	0.75	0.66	0.75	0.07	0.39	0.14	0.31
Is forgetful in daily activities	-0.09	0.28	0.73	0.58	0.73	0.58	-0.05	0.29	-0.05	0.23
Hyperactivity										
Fidgets with hands or feet or squirms in seat	0.56	0.69	0.19	0.37	0.14	0.38	0.58	0.70	0.07	0.42
Leaves seat in classroom	0.56	0.65	0.17	0.20	0.08	0.21	0.71	0.66	-0.06	0.23
Has difficulty playing quietly	0.59	0.72	0.20	0.26	0.13	0.27	0.66	0.73	0.03	0.30
Runs about or climbs excessively	0.65	0.67	0.07	0.21	-0.02	0.22	0.79	0.68	-0.03	0.24
Often sense of an extreme internal restlessness	0.45	0.49	0.09	0.20	0.03	0.21	0.52	0.50	0.00	0.23
Permanently extremely restless	0.69	0.67	0.04	0.23	-0.04	0.24	0.77	0.68	0.03	0.26
Is "on the go" or acts as if driven by a motor	0.56	0.46	-0.05	0.21	-0.10	0.22	0.57	0.46	0.08	0.24
Impulsivity										
Blurts out answers	0.64	0.53	-0.07	0.24	0.04	0.23	-0.07	0.31	0.79	0.60
Has difficulty awaiting turn	0.69	0.62	-0.01	0.28	0.05	0.27	0.15	0.37	0.63	0.71
Interrupts or intrudes on others	0.65	0.64	0.06	0.28	0.11	0.26	0.17	0.36	0.57	0.73
Talks excessively	0.67	0.45	-0.17	0.24	-0.07	0.23	0.00	0.31	0.76	0.52
Eigenvalue unrotated	7.27		1.92		7.27		1.92		1.05	
Variance explained (PCA)	45.95				51.21					
RMSEA (CFA)	0.07				0.06					
AGFI (CFA)	0.98				0.99					

Comparing the two- and the three-component solutions

Table 2 Pattern coefficients: Confirmatory factor analysis using structural equation modeling of the Conners' Index item correlation matrix

Items ^a	Loading	Goodness of fit	
		RMSEA	AGFI
Restless or overactive	0.63	0.18	0.96
Excitable, impulsive	0.72		
Disturbs other children	0.70		
Fails to finish things he starts	0.57		
Fidgeting	0.76		
Inattentive, easily distracted	0.69		
Demands must be met immediately	0.59		
Cries easily	0.45		
Mood changes quickly and drastically	0.62		
Temper outbursts, explosive or unpredictable behaviour	0.69		

^aItems represent the German translation; some slight deviation from the original wording was required to achieve conceptual equivalence with the original instrument

tern coefficients ranged between 0.45 and 0.78. The three-component structure resembled the theoretical measurement model with the dimensions inattention, hyperactivity, and impulsivity. A total of 51.2% of the variance was explained. The pattern coefficients ranged between 0.52 and 0.81 (see Table 1).

A confirmatory factor analysis was conducted by specifying a structural equation model according to the 2-dimensional (D) measurement model, incorporating the latent variables inattention and hyperactivity/impulsivity. Table 1 shows the results and informs about the structure of the model. For the entire sample of children and adolescents aged 7–17 years, the confirmatory factor analysis of the parent-reported FBB-HKS resulted in goodness of fit (GoF) statistics of RMSEA = 0.07, indicating a good fit of the 2-D measurement model. The AGFI was 0.98, indicating that 98% of the observed variance and covariance could be explained by the model. The estimated loading coefficients ranged between 0.45 and 0.75. The largest cross loading was 0.40. None of the cross loadings exceeded the estimated loading on the domain the item was intended to measure. The estimated latent constructs inattention and hyperactivity/impulsivity correlated with $r = 0.70$.

Next, a structural equation model was specified according to the 3-dimensional (D) measurement model incorporating the latent variables inattention, hyperactivity, and impulsivity. Table 1 shows the results and informs about the structure of the model. For the entire sample of children and adolescents aged 7–17 years, this confirmatory factor analysis resulted in GoF statistics of RMSEA = 0.06, indicating a good fit of the 3-D model. The AGFI was 0.99. The estimated loading coefficients ranged between 0.46 and 0.75. The

largest cross loading was 0.41. None of the cross loadings exceeded the estimated loading on the domain the item was intended to measure. The estimated latent construct inattention correlates with hyperactivity $r = 0.72$ and impulsivity $r = 0.57$; hyperactivity and impulsivity scores correlate with $r = 0.79$.

A multi-group analysis was conducted to test for statistically significant differences in the pattern coefficients estimated for younger (7–10 years) and older (11–17 years) respondents (data not shown). For the 2-D model, the GoF Chi-squared value was 6334.36 ($df = 378$) for the model with pattern coefficients restricted to be equal across groups. For an unrestricted model with separate estimation of pattern coefficients, the GoF Chi-squared value was 6136.21 ($df = 338$). The resulting difference in Chi-squared values of 198.15 ($df = 40$) was $P < 0.001$; thus, indicating statistically significant differences in the pattern coefficient estimates for younger and older respondents. For the 3-D model, the GoF Chi-squared value was 5380.61 ($df = 374$) for the model with pattern coefficients restricted to be equal across groups. For an unrestricted model with separate estimation of pattern coefficients, the GoF Chi-squared value was 5178.64 ($df = 334$). The resulting difference in Chi-squared values of 201.97 ($df = 40$) was $P < 0.001$, indicating statistically significant differences in the pattern coefficient estimates. Examining the actual loading coefficients issued from separate estimation for children aged 7–10 versus 11–17 years showed similar coefficients across age for the 2- as well as for the 3-D model, except for one item: For the item “Is ‘on the go’...”, the estimated loading coefficients were 0.68 for the younger and 0.53 for the older respondents in both the 2- and the 3-D model.

The unidimensionality of the Conners' Index was tested next (Table 2). A one-factorial structural equation model was specified and tested. The loading coefficients issued from the confirmatory factor analysis of the Conners' scale ranged between 0.45 and 0.76. The one-factorial model of the Conners' Index could not adequately explain the pattern of correlation between the items. The RMSEA was 0.18; thus, indicating a poor GoF. However, the AGFI of 0.96 was acceptable. The residual correlation between item 1 and 5 ($r_{res} = 0.23$) as well as between item 4 and 6 ($r_{res} = 0.21$) were slightly larger than the threshold used by Bjorner et al. [6] as an indicator for violation of the unidimensionality assumption in a set of items.

A multi-group analysis was again conducted to test for statistical significant differences in the pattern coefficients estimated for younger (7–10 years) and older (11–17 years) children and adolescents. The GoF Chi-squared value was 3585.17 ($df = 90$) for the model with pattern coefficients restricted to be equal

Table 3 Internal consistency (Cronbach's α), means (SDs), missing values, and ceiling and floor effect for the total sample

Scales	7–10 years (<i>n</i> = 960)		11–13 years (<i>n</i> = 750)		14–17 years (<i>n</i> = 1,153)		Effect size		ALL (<i>n</i> = 2,863)		At ceiling (%) At floor (%)	Missing values (%) ^a
	α	Mean (SD)	α	Mean (SD)	α	Mean (SD)	SMD 7–10 vs. 11–13	SMD 7–10 to 14–17	α	Mean (SD)		
FBB-HKS												
Inattention ^{b,c}	0.88	0.72 (0.51)	0.88	0.65 (0.48)	0.88	0.63 (0.48)	0.13	0.19	0.88	0.67 (0.48)	6.4 0.0	0.3
Hyperactivity ^c	0.83	0.34 (0.44)	0.78	0.26 (0.35)	0.76	0.26 (0.34)	0.21	0.21	0.79	0.29 (0.38)	32.9 0.0	0.9
Impulsivity ^c	0.70	0.60 (0.51)	0.71	0.45 (0.48)	0.74	0.42 (0.49)	0.30	0.38	0.73	0.49 (0.50)	29.1 0.0	0.6
Hyperactivity ^b	0.86	1.44 (0.42)	0.84	1.33 (0.36)	0.83	1.32 (0.35)	0.27	0.31	0.84	1.36 (0.38)	16.8 0.0	0.9
Impulsivity ^b											0.0	
Total	0.91	0.56 (0.41)	0.90	0.48 (0.37)	0.89	0.46 (0.37)	0.23	0.27	0.90	0.50 (0.39)	3.6 0.0	0.9
Conners' Index	0.85	0.75 (0.55)	0.84	0.71 (0.52)	0.82	0.58 (0.48)	0.08	0.33	0.84	0.67 (0.52)	5.2 <0.1	1.2

^aparent's questionnaire was returned
^b2 - factor solution
^c3 - factor solution

across groups. For an unrestricted model with separate estimation of pattern coefficients, the GoF Chi-squared value was 3488.18 ($df = 70$). The resulting difference in Chi-squared values of 96.99 ($df = 20$) was $P < 0.001$; thus, indicating statistically significant differences in the pattern coefficients across age. The largest differences in loading coefficients was 0.64 (7–10 years) versus 0.54 (11–17 years) for the item “Fails to finish things he/she starts”.

■ Reliability and psychometric properties

For the FBB-HKS total score, the internal consistency was Cronbach's $\alpha = 0.90$. Subscale scores ranged from 0.73 to 0.88. Table 3 shows only slight differences in α across age groups. On average, younger respondents achieved slightly higher scores on all FBB-HKS scales (more symptoms). The magnitude of this effect was up to 0.38 of a standard deviation between 7 to 10 and 14 to 17-year-olds on the impulsivity scale. Some ceiling effects were observed for the scores of the 3-D FBB-HKS measurement model. For the Conners' Index, Cronbach's α was 0.84 and only slightly differed across age groups. Again, younger respondents scored slightly higher. The magnitude of the effect was 0.33 of a standard deviation between 7 to 10 and 14 to 17-year-olds.

Boys displayed higher scores than girls on all FBB-HKS scores as well as on the Conners' Index. This effect was largest for the FBB-HKS inattention scale. Low socioeconomic status was also associated with higher scores on all scales of the FBB-HKS and on the Conners' Index. The largest effect was found for the FBB-HKS total score (Table 4).

Among the different scoring opportunities of the FBB-HKS, the total score displayed the highest correlation with the SDQ scales. As expected a priori, the correlation was highest with the SDQ hyperactivity scale ($r = 0.69$). Except for the impulsivity score, all other FBB-HKS scores also correlated highest with SDQ hyperactivity. The Conners' Index correlated highest ($r = 0.66$) with the SDQ total difficulties score. The correlation with SDQ hyperactivity was $r = 0.62$ and very close to the correlation with SDQ conduct problems. An inspection of the item content reveals that several items of the Conners' Index describe oppositional behaviour problems or emotional problems that are not core symptoms of ADHD as defined by the ICD-10 and DSM-IV. Correlations between FBB-HKS scores and the Conners' Index ranged from 0.55 (Conners with Impulsivity) to 0.70 (Conners with FBB-HKS total) (Table 5).

Discussion

This paper examined the psychometric properties of two ADHD screening instruments, the FBB-HKS and the Conners' Index. The Conners' Index provides a unidimensional index of ADHD symptoms and burden while the FBB-HKS allowed for scoring in three different ways. In addition to a unidimensional index of overall ADHD symptoms and burden (scoring A), a 2-D model of inattention and hyperactivity/impulsivity is also possible (scoring B). A third variant (scoring C) measures the 3-D of inattention, hyperactivity, and impulsivity. Different scoring alternatives were examined in this paper; however, making a

Table 4 Differences in mean (SD) for gender and low, medium, and high socioeconomic status (SES) (Winkler Index)

Scales	Boys (n = 1467)		Girls (n = 1396)		Effect size gender		Low SES (n = 716)		Medium SES (n = 1330)		High SES (n = 789)		Effect size high vs. medium		Effect size high vs. low	
	Mean	(SD)	Mean	(SD)	SMD	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	SMD	(SD)	SMD	(SD)
FBB-HKS																
Inattention ^{a,b}	0.76	0.52	0.56	0.44	0.41	0.54	0.77	0.50	0.66	0.50	0.59	0.43	0.21	0.36		
Hyperactivity ^b	0.35	0.43	0.23	0.31	0.30	0.45	0.38	0.37	0.28	0.37	0.24	0.32	0.26	0.38		
Impulsivity ^b	0.52	0.53	0.46	0.47	0.11	0.59	0.62	0.47	0.48	0.47	0.41	0.45	0.27	0.41		
Hyperactivity ^a	1.41	0.42	1.32	0.33	0.25	0.46	1.47	0.37	1.35	0.37	1.30	0.32	0.30	0.44		
Impulsivity																
Total	0.57	0.42	0.43	0.33	0.37	0.45	0.60	0.38	0.49	0.38	0.43	0.32	0.28	0.45		
Conners' Index	0.73	0.55	0.61	0.48	0.23	0.59	0.77	0.51	0.67	0.51	0.59	0.46	0.19	0.33		

^a2 - factor solution^b3 - factor solution

suggestion in favour of either instrument or scoring alternative is not the main goal. Instead, we aim to discuss the pros and cons of the alternatives and in this way, help to decide for which purpose and under which circumstances the different alternatives should be applied.

The factorial validity of the different FBB-HKS scorings could be confirmed in confirmatory factor analysis. The models specified according to scoring B and C fit the data well, accounting for the empirical pattern of inter-item correlation in an adequate way. The more complex scoring C with its 3-D was only slightly superior to the more parsimony scoring B (2-D). The empirical justification of the unidimensional global index (scoring A) was not directly tested. However, the large correlations between the components inattention, hyperactivity, and impulsivity (examined in the confirmatory factor analysis as well as in the inter-scale correlation) hint at the justification of an overall index. Assuming that the inter-scale correlation is caused by an underlying general factor, the assessment of such a general factor would be justified. We did not specify and test second order models incorporating such a general factor because the actual models with correlated latent factors led to similar goodness of fit values. Nevertheless, one must bear in mind that the usage of the global scoring could lead to a loss of psychometric information, pointing into different directions [18]. Given the good fit of the 2- and 3-D models, we do not consider the global index as being strictly unidimensional.

The same applies to the unidimensional Conners' Index. The confirmatory factor analysis shows that this index is not strictly unidimensional. Psychometric information pointing into different directions is combined into one single value and, thus, might be lost. However, the actual results hint at only a slight deviation from the unidimensionality assumption.

An examination of the factorial stability across age groups revealed deviation from factorial invariance in the FBB-HKS scorings, which could be attributable to the item "Is 'on the go'...". This item seems to contribute less to the measurement for older than for younger respondents. Nevertheless the validity of the measurement model itself across age groups can be assumed since the actual loading coefficients for the older respondents still indicate a substantial contribution of the item to the measurement. A similar result was seen for the item "Does not finish tasks" of the Conners' Index, which contributes less to the assessment for older children. Future research employing analyses rooted in the item response theory could focus on explicitly testing for differential item functioning (DIF) [11]. Based on such research, techniques of adjusting for DIF could be developed and applied.

Table 5 Correlations (Pearson) between FBB-HKS and Conners' Index and the strengths and difficulties questionnaire (SDQ) symptom scales

	FBB-HKS					Conners' Index	SDQ				
	Inattention	Hyperactivity	Impulsivity	Hyperactivity Impulsivity	Total	Total	Total Difficulties Score	Emotional Symptoms	Conduct Problems	Hyperactivity	Peer Problems
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
FBB-HKS											
Inattention ^{a,b}		0.59	0.45	0.59	0.60	0.63	0.35	0.45	0.65	0.26	
Hyperactivity ^b			0.61		0.61	0.53	0.26	0.41	0.57	0.20	
Impulsivity ^b					0.55	0.45	0.19	0.41	0.44	0.20	
Hyperactivity ^a		0.92	0.86		0.65	0.55	0.26	0.46	0.57	0.22	
Impulsivity											
Total	0.90	0.85	0.73	0.89	0.70	0.66	0.34	0.50	0.69	0.27	
Conners' Index						0.66	0.38	0.56	0.62	0.28	

All correlations significant at the 0.01 level (2-tailed)

^a2 - factor solution

^b3 - factor solution

The examination of internal consistency of item responses showed satisfying results for the different scorings of the FBB-HKS as well as for the Conners' Index. However, only the global scoring of the FBB-HKS achieved the reliability required for individual comparisons or for an individual assessment (e.g., monitoring treatment) as demanded by Nunnally and Bernstein [36]. It must be noted, however, that other authors have proposed more liberal criteria (e.g., Lienert and Raatz [33]).

The observed differences between gender, age groups, and socioeconomic status groups are consistent with results found in other studies [3, 19] as well as with the findings by Döpfner et al. and Huss et al. in this supplement [20, 30]. The FBB-HKS displayed slightly larger sensibility for these differences. The pattern of correlation with the scales of the strengths and difficulties questionnaire hinted at convergent validity of the FBB-HKS scorings and the Conners' Index. However, the discrimination between ADHD and conduct problems is somewhat better for the FBB-HKS than for the Conners' Index.

The results obtained with the FBB-HKS replicated earlier findings in smaller, representative German samples regarding factor structure, reliability, and

convergent / discriminant validity [9, 10, 23, 26] as well as studies with similar DSM-IV-based rating scales from other cultures [21, 24, 43]. To summarise, the two instruments provide a reasonable assessment of ADHD symptoms and burden with good psychometric properties.

The decision for only one instrument should be based on the actual aims of the study. In general, the larger FBB-HKS was slightly superior to the Conners' Index. Since the items of the FBB-HKS are closely linked to the symptom criteria of the DSM-IV and ICD-10, the content validity of this scale regarding ADHD as defined by the classification schemes may also be higher.

This analysis has been conducted in a representative sample of children and adolescents. Thus, the results largely represent psychometric characteristics of the scales in assessing the normal variation of ADHD symptoms. Additional analyses in clinical samples of children with elevated ADHD scores are necessary in order to analyse the psychometric quality of these scales in the clinical range.

■ **Conflict of interest** All authors declare no conflict of interest.

References

1. Achenbach TM, Edelbrock C (1983) Manual for the Child Behavior Checklist and Revised Child Behavior Profile. University of Vermont, Department of Psychiatry, Burlington
2. Arnold LE, Barnebey NS, Smeltzer DJ (1981) First grade norms, factor analysis and cross correlation for Conners, Davids, and Quay-Peterson behavior rating scales. *J Learn Disabil* 14:269–275
3. Becker A, Rothenberger A, Deutsche ADORE/FACE Studiengruppe (2006) Psychopathological Screening of children with ADHD: strengths and difficulties questionnaire in a paneuropean study. *Eur Child Adolesc Psychiatry* 15:56–62
4. Biederman J (2005) Attention-deficit/hyperactivity disorder: a selective overview. *Biological Psychiatry* 57:1215–1220
5. Birnbaum HG, Kessler RC, Lowe SW, Secnik K, Greenberg PE, Leong SA, Swensen AR (2005) Costs of attention deficit-hyperactivity disorder (ADHD)

- in the US: excess costs of persons with ADHD and their family members in 2000. *Curr Med Res Opin* 21:195–206
6. Bjorner JB, Kosinski M, Ware JE (2003) Calibration of an item pool for assessing the burden of headaches: An application of item response theory to the Headache Impact Test (HIT™). *Qual Life Res* 12:913–933
 7. Blank R, Remschmidt H (1992) Subgruppen hyperkinetischer Störungen-explorative Untersuchungen unter Berücksichtigung von Fragebogenverfahren und immunologischen Parametern. Subgroups of hyperkinetic disorders—explorative studies with reference to questionnaires and immunologic parameters. *Z Kinder Jugendpsychiatr* 20:34–45
 8. Breuer D, Döpfner M (2006) Aufmerksamkeitsdefizit- / Hyperaktivitätsstörungen bei Drei- bis Sechsjährigen in der ärztlichen Praxis—eine bundesweite Befragung. *Z Kinder Jugendpsychiatr Psychother* 34:357–365
 9. Breuer D, Döpfner M (2008) Validierung eines Fragebogens zur Erfassung von Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen (ADHS) bei Vorschulkindern im Eltern- und im Erzieherurteil. *Z Entwicklungspsychol Pädagog Psychol* (accepted)
 10. Brühl B, Döpfner M, Lehmkuhl G (2000) Der Fremdbeurteilungsbogen für hyperkinetische Störungen (FBB-HKS)—Prävalenz hyperkinetischer Störungen im Elternurteil und psychometrische Kriterien. *Kindheit und Entwicklung* 9:116–126
 11. Camilli G, Shepard LA (1994) *Methods for identifying biased test items*. Sage Publications, London
 12. Campbell SB, Steinert Y (1978) Comparisons of rating scales of child psychopathology in clinic and nonclinic samples. *J Consult Clin Psychol* 46:358–359
 13. Coghill D, Spiel G, Baldrsson G, Döpfner M, Lorenzo MJ, Ralston SJ, Rothenberger A, the ADORE study group (2006) Which factors impact on clinician-rated impairment in children with ADHD? *Eur Child Adolesc Psychiatry* 15(Suppl 1):30–37
 14. Collett BR, Ohan JL, Myers KM (2003) Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 42:1015–1037
 15. Conners C (1970) Symptom patterns in hyperkinetic, neurotic, and normal children. *Child Dev* 41:667–682
 16. Conners CK (1973) Rating scales for use in drug studies with children. *Pharmacotherapy of children [Special issue]*. *Psychopharmacol Bull*, 24–84
 17. Conners CK, Sitarenios G, Parker JD, Epstein JN (1998) The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *J Abnorm Child Psychol* 26(4):257–268
 18. Coste J, Bouée S, Ecosse E, Leplège A, Pouchot J (2005) Methodological issues in determining the dimensionality of composite health measures using principal component analysis: case illustration and suggestions for practice. *Qual Life Res* 14:641–654
 19. Cuffe SP, Moore CG, McKeown RE (2005) Prevalence and correlates of ADHD symptoms in the National Health Interview Survey. *J Atten Disord* 9:392–401
 20. Döpfner M, Breuer D, Wille N, Erhart M, Ravens-Sieberer U, BELLA study group: How often do children meet ICD-10/DSM-IV criteria of attention deficit-/hyperactivity disorder and hyperkinetic disorder? Parent-based prevalence rates in a national sample – results of the BELLA study. *Eur Child Adolesc Psychiatry* 17(Suppl1):59–70
 21. Döpfner M, Steinhausen H-C, Coghill D, Dalsgaard S, Poole L, Ralston SJ, Rothenberger A, the ADORE Study Group (2006) Cross-cultural reliability and validity of ADHD assessed by the ADHD Rating Scale in a pan-European study. *Eur Child Adolesc Psychiatry* 15:S46–S55
 22. Döpfner M, Lehmkuhl G (2000) Diagnostik-System für Psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV (DISYPS-KJ), 2nd edn. Huber, Bern
 23. Döpfner M, Görtz-Dorten A, Lehmkuhl G (2008) Diagnostik-System für Psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV, DISYPS-II. Huber, Bern
 24. DuPaul GJ, Power T, Anastopoulos AD (1998) *ADHD Rating Scales-IV: checklist, norms, and clinical interpretation*. Guilford, New York
 25. Goodman R (1997) The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatr Allied Discip* 38:581–586
 26. Görtz A, Döpfner M, Nowak A, Bonus B, Lehmkuhl G (2002) Ist das Selbsturteil Jugendlicher bei der Diagnostik von Aufmerksamkeitsdefizit- / Hyperaktivitätsstörungen hilfreich? Eine Analyse mit dem Diagnostiksystem DISYPS. *Kindheit und Entwicklung* 11:82–89
 27. Görtz-Dorten A, Döpfner M (2008) Aufmerksamkeitsdefizit- / Hyperaktivitätsstörungen von Jugendlichen im Selbst- und im Elternurteil –eine Analyse an einer Feldstichprobe mit dem Diagnostiksystem DISYPS-II. *Z Kinder Jugendpsychiatr Psychother* (under review)
 28. Hechtman L (2000) Assessment and diagnosis of attention-deficit/hyperactivity disorder. *Child Adolesc Psychiatry Clin North Am* 9:481–498
 29. Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Struct Equation Modeling* 6:1–55
 30. Huss M, Hölling H, Kurth B-M, Schlack R (2008) How often are German children and adolescents diagnosed with ADHD? Prevalence based on the judgment of health care professionals: results of the German Health and Examination Survey (KiGGS). *Eur Child Adolesc Psychiatry* 17(Suppl1): 52–58
 31. Jöreskog KG, Sörbom D (2001) *LISREL 8: Users's Reference Guide*. Scientific Software International, Inc., Lincolnwood
 32. Jöreskog KG, Sörbom D (2002) *PRELIS 2: Users's Reference Guide*. Scientific Software International, Inc., Lincolnwood
 33. Lienert GA, Raatz U (1998) *Testaufbau und Testanalyse*, 6th edn. Beltz, Weinheim
 34. Mash EJ, Johnston C (1983) Parental perceptions of child behavior problems, parenting self-esteem, and mothers' reported stress in younger and older hyperactive and normal children. *J Consult Clin Psychol* 51:6899
 35. McDonald RP, Ho MHR (2002) Principals and Practice in reporting structural equation analyses. *Psychol Methods* 7:64–82
 36. Nunnally JC, Bernstein IR (1994) *Psychometric Theory*, 3rd edn. McGraw-Hill, New York
 37. Ravens-Sieberer U, Kurth B-M, KiGGS study group, BELLA study group (2008) The mental health module (BELLA study) within the German Health Interview and Examination Survey of Children and Adolescents (KiGGS): study design and methods. *Eur Child Adolesc Psychiatry* 17(Suppl1): 10–21
 38. Riley AW, Coghill D, Forrest CB, Lorenzo MJ, Ralston SJ, Spiel G, the ADORE study group (2006) Validity of the health-related quality of life assessment in the ADORE study: Parent report form of the CHIP-CE edition. *Eur Child Adolesc Psychiatry* 15:S63–S71
 39. Starfield B, Forrest CB, Ryan SA, Riley AW, Ensminger ME, Green BF (1996) Health status of well vs. ill adolescents. *Arch Pediatr Adolesc Med* 150:1249–1256

40. Swanson J (1992) School based assessments and interventions for ADD students. KC Publishing, Irvine
41. Ullmann RK, Sleator EK, Sprague RL (1985) A change of mind: the Conners abbreviated rating scales reconsidered. *J Abnorm Child Psychol* 13:553–565
42. Winkler J, Stolzenberg H (1999) Social status scaling in the German National Health Interview and Examination Survey. *Gesundheitswesen* 61:S178–S183
43. Wolraich ML, Lambert W, Doffing MA, Bickman L, Simmons T, Worley K (2003) Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *J Pediatr Psychol* 28:559–568