Title: An 8-item short form of the *Eating Disorder Examination-Questionnaire* adapted for children (*ChEDE-Q8*).

Authors: Sören Kliem[a], Ricarda Schmidt[b,c], Mandy Vogel[d,e], Andreas Hiemisch[d,e], Wieland Kiess[d,e], & Anja Hilbert[b,c]

Author Affiliations:

[a]Criminological Research Institute of Lower Saxony, Hannover, Germany

[b]Integrated Research and Treatment Center Adiposity Diseases, University of Leipzig Medical Center, Germany

[c]Department of Medical Psychology and Medical Sociology, University of Leipzig, Germany

[d]Hospital for Children and Adolescents, Center for Pediatric Research, University of Leipzig Medical Center, Germany

[e]LIFE Leipzig Research Center for Civilization Diseases, University of Leipzig Medical Center, Germany

Corresponding Author:

Dr. Sören Kliem, Dipl.-Psych.

Criminological Research Institute of Lower Saxony

Lützerodestrasse 9

30161 Hannover, Germany

Phone: +49 (0) 511 34836-70

Fax: +49 (0) 511 34836-10

E-mail: Soeren.Kliem@kfn.de

Abstract

Objective: Eating disturbances are common in children placing a vulnerable group of them at risk for full-syndrome eating disorders and adverse health outcomes. In order to provide a valid self-report assessment of eating disorder psychopathology in children, a short form of the child version of the *Eating Disorder Examination* (*ChEDE-Q*) was psychometrically evaluated. Similar to the *EDE-Q*, the *ChEDE-Q* provides assessment of eating disorder psychopathology related to anorexia nervosa, bulimia nervosa, and binge-eating disorder. However, it does not assess symptoms of avoidant/restrictive food intake disorder, pica, or rumination disorder.

Method: In 1836 participants ages 7 to 18 years, recruited from two independent population-based samples, the factor structure of the recently established 8-item short form *EDE-Q8* for adults was examined, including measurement invariance analyses on age, gender, and weight status derived from objectively measured weight and height. For convergent validity, the *ChEDE-Q* global score, Body Esteem Scale, Strengths and Difficulties Questionnaire, and sociodemographic characteristics were used. Item characteristics and age- and gender-specific norms were calculated.

Results: Confirmatory factor analysis revealed good model fit for the 8-item *ChEDE-Q*. Measurement invariance analyses indicated strict invariance for all analyzed subgroups. Convergent validity was provided through associations with well-established questionnaires and age, gender, and weight status, in expected directions.

Discussion: The newly developed *ChEDE-Q8* proved to be a psychometrically sound and economical self-report assessment tool of eating disorder psychopathology in children. Further validation studies are needed, particularly concerning discriminant and predictive validity.

An 8-item short form of the Eating Disorder Examination-Questionnaire adapted for

children (*ChEDE-Q8*)

Eating disturbances are relatively common in childhood and adolescence,[1–7] including a wide range of eating or weight-related symptoms (i.e., loss of control eating, binge eating, avoidant or restrictive eating), compensatory strategies of weight control (i.e., self-induced vomiting or misuse of laxatives), and associated general eating disorder psychopathology, including dietary restraint and concerns about body image and weight.[8] Although the tendency to remit is high in children[9], an early manifestation of eating disturbance is a frequently reported risk factor for later full-blown eating disorders,[10–14] increased psychopathology,[12,15–18] obesity,[12,15,16,18] and impairment in quality of life[19].

Given the lack of psychometrically sound instruments for children and youth, the child version of the *Eating Disorder Examination-Questionnaire (EDE-Q)*,[20] the *ChEDE-Q,* was developed for children aged 8-14 years[21], and evaluated in samples of children and adolescents aged 10-16[22], 12-17[23] and 8-13[24] years. This self-report instrument contains, like the adult version, 22 items allocated to four subscales (restraint, eating concern, weight concern, shape concern) which can be integrated into one global score of eating disorder psychopathology. Six additional key behavioral items measure diagnostically relevant information, for example, binge eating or self-induced vomiting.

With the *EDE-Q8*[25] an 8-item short form for the assessment of eating disorder psychopathology was developed and evaluated using two representative population surveys in Germany, for use in adults and adolescents aged ≥ 14 years. The *EDE-Q8* had excellent item characteristics (r_{it} >.30), very good reliability (α>.90), and a very good model fit for the postulated second-order factorial structure. Furthermore, a strong correlation (*r*=.75) between the *EDE-Q8* and a 13-item short form of the Eating Attitudes Test[26] was observed.

Along the lines of the *EDE-Q*, the *ChEDE-Q* provides assessment of eating disorder psychopathology related to anorexia nervosa (AN), bulimia nervosa (BN), and binge-eating

disorder (BED), but it does not assess symptoms of avoidant/restrictive food intake disorder (ARFID), pica, or rumination disorder. The aim of this study was to evaluate a comparable short form of the *ChEDE-Q,* based on an item selection of the *EDE-Q8*, for an economical assessment of eating disorder psychopathology related to AN, BN, and BED in children and youth.

Method

Sampling Procedure

Children and adolescents were recruited for two different research projects at the Universities of Marburg (Sample 1) and Leipzig (Sample 2): Sample 1 included an unselected community-based sample, recruited for a research project on loss of control eating between 2005-2007 (for methodological detail see Hilbert and Czaja[27]). *N*=594 8-13 year old children were assessed with the *ChEDE-Q*. Sample 2 is part of the ongoing 'Leipzig Research Centre for Civilization Diseases (LIFE)' Child study which started in 2011. One aim of this prospective population-based cohort study is to identify risk factors of childhood obesity and associated mental disorders (for methodological detail see Quante et al.[28]). *N*=1242 7-18 year old youth were investigated with the *ChEDE-Q*. Broadening of the age range was allowed to evaluate suitability for higher (and lower) ages than postulated.[20] Using multigroup confirmatory factor analysis, we found strict measurement invariance (see Statistical Analyses) between both samples (see Appendix A). Thus, pooling both samples was considered to be adequate.

Participants

Sample characteristics of the total sample (*N*=1836) are displayed in Table 1. The two projects were approved by the Ethics Committees of the German Psychological Society and the Medical Faculty of the University of Leipzig, Germany (Reg. No. 264-10-19042010). Informed consent and assent were obtained from at least one parent of the under-aged participants.

Measures

In sample 2, the *appearance* subscale of the child version of the *Body Esteem Scale* (BES)[29], containing 12-items on a 4-point Likert scale (1 = "strongly disagree" to 4 = "strongly agree"), and the *emotional problems* subscale of the *Strengths and Difficulties Questionnaire* (SDQ)[30], containing five items on a 3-point scale (0 = "not true" to 2 ="certainly true") were included. Higher scores on the *BES* indicate greater body satisfaction; higher scores on the *SDQ* subscale indicate more emotional problems. The observed internal consistencies of the subscales of the *BES* (α=.92) and *SDQ* (α=.69) were excellent to acceptable. The body mass index (BMI, kg/m²) of children was calculated from objectively measured weight and height in sample 2 (height and weight were not measured for sample 1) and subsequently standardized regarding age and sex using German reference data (i.e., BMI standard deviation scores [SDS]).[31]

Statistical Analyses

The German version of the *ChEDE-Q8* was used, embedded in the *ChEDE-Q*[32] (i.e., children completed the *ChEDE-Q*, but only eight items were considered in the data analyses; for items see Table 2). At the item level, means and standard deviations were determined. Item selectivity indicating if one item varies in line with other items was computed as the correlation of the respective item with the sum of all other items (r_{it}, item-rest correlation coefficients). In accordance with Field[33], item selectivity values less than 0.30 can be classified as unsatisfactory. Item difficulty coefficients were calculated as quotients of the sum of the item values that were obtained, and the sum of the maximum achievable item values, multiplied by 100 (possible range: 0 [none of the participants answered the item positively] to 100 [all of the participants answered the item positively]). Differences in *ChEDE-Q8* global scores were analyzed using two-factorial (age and gender) analysis of variance (ANOVA). Missing data (0.1–0.7% per item) were imputed using chained equation

modeling,[34] based on gender and age. Predictive mean matching was used for imputation (i.e.,

only realistic values were computed). Construct validity of the *ChEDE-Q8* was examined

through positive correlations with the full *ChEDE-Q* (including and excluding *ChEDE-Q8*

items), the *emotional problems* subscale of the *SDQ*, and weight (i.e., BMI-SDS).

Furthermore, we expected a negative correlation with the *appearance* subscale of the child

version of the *BES*. Based on the *EDE-Q8*,[25] factorial validity was tested using confirmatory

factor analysis (CFA) by a second-order general factor model with four first-order factors

comprising the postulated subscales. We used robust maximum likelihood estimation with the

mean-adjusted Satorra-Bentler χ^2 test statistic.[35] To evaluate goodness-of-fit, the following

four criteria were considered: The Standardized Root Mean Square Residual (SRMR) and the

Root Mean Square of Approximation (RMSEA) with its 90% confidence interval (CI) were

used to assess absolute model fit. The Comparative Fit Index (CFI) and Tucker Lewis Index

(TLI) measure relative model fit compared to the "null" model. RMSEA and SRMR values <

.050 represent a close fit, values between .050 and .080 a reasonably close fit, and values >

.080 an unacceptable model.[30] Regarding CFI and TLI, Hu and Bentler[36] suggested a CFI and

TLI >.900 for an adequate fit and a CFI and TLI >.950 for a good model fit.

The *ChEDE-Q8* is a complex instrument and it could be considered challenging to

complete for children at the younger end of the age range included in this study. For this

reason, we investigated age-related effects on the measurement model of the *ChEDE-Q8*

using analysis of measurement invariance (group 1: 7 to 10 years; group 2: >10 to 12 years;

group 3: >12 to 14 years; group 4: >14 to 18 years, see Table 1). Furthermore, we conducted

measurement invariance tests across gender and weight status (obesity and overweight vs.

normal weight). Underweight was not analyzed in measurement invariance testing because of

few underweight cases

In accordance with the sequential strategy developed by Meredith and Teresi[37], the

weak invariance model was first tested by constraining the estimate factor loadings to be

equal across groups. If empirical support for weak invariance is provided, it allows for the comparison of structural relationships between latent constructs in groups. Second, the strong invariance model was tested by constraining both loadings and intercepts to be equal across groups. This level of invariance allows for the comparison of means of the latent construct between groups. Finally, we tested the strict invariance model by constraining the loadings, intercepts, and item error variances to be equal across groups. Lack of strict invariance can affect decisions in screening processes that depend on the expression of a construct, resulting in different error rates (i.e., sensitivity, specificity) for different groups.[38]

As recommended by Chen,[39] a change of $\Delta\text{CFI} \leq -.010$ in CFI, supplemented by a change of $\Delta\text{RMSEA} \geq .015$, was considered as indicative of non-invariance. Data analysis was carried out using the R packages lavaan[40] and mice.[41]

Results

Item Characteristics

Means and standard deviations, and item characteristics of the *ChEDE-Q8* are depicted in Table 2. The global mean score of the *ChEDE-Q8* was 1.05 (*SD*=1.33), with an observed range from 0 to 6. Item difficulty values (p_i) ranged between p_i=9 (preoccupation with food) and p_i=27 (desire to lose weight). Regarding item selectivity, all item-rest correlation coefficients exceeded the relevant cut-off (range: r_{it}=.43 [preoccupation with food] to r_{it}=.80 [feelings of fatness]). For some items, there were statistically significant differences between boys and girls, however, with negligible effect sizes (d=0.04 to 0.12).

Norms

A two-factorial ANOVA with the factors gender and age resulted in a significant interaction effect (F(3, 1836)=6.50; p<.001). Thus, age- and gender-specific norms are provided in Table 3. For practical reasons, we determined severity categories in accordance with the *ChEDE-Q8* percentile ranks. The level of eating disorder psychopathology was

classified as critical if the percentiles were between the 90th and 95th percentile, as high if the percentiles were between the 95th and 98th, and as very high if the percentiles were above the 98th percentile.

Internal Consistency

With regard to the global score of the *ChEDE-Q8*, the internal consistency for the total sample was α=.89 (boys: α=.87; girls: α=.90).

Factorial Validity

CFA revealed good fit parameters for the second-order general factor model (SRMR=.042; RMSEA=.077, 90% CI [.070; .083]; CFI=.946; TLI=.916). Second-order corrected factor loadings were medium to high (.41 to .87). Detailed results are displayed in Appendix B. Measurement invariance analyses indicated strict invariance for all of the analyzed subgroups. Detailed results can be found in Appendix A of the supporting information.

Construct Validity

Correlations of the *ChEDE-Q8* global score with the *ChEDE-Q* global score (r=.97, p<.001) as well as with the *ChEDE-Q* global score not including the items of the short form (r=.91; p<.001) were very high. The correlation coefficients were as follows: *BES*: r=-.65, p<.001; *SDQ emotional problems*: r=.37, p<.001 BMI-SDS: r=.51, p<.001.

Discussion

Based on a large-scale population-based sample of more than 1800 children, the present study provided a psychometrically sound 8-item short form of the commonly used *ChEDE-Q* for assessing children's eating disorder psychopathology. The newly developed *ChEDE-Q8* showed good internal consistency, a clear second-order one-factorial structure, and satisfactory convergent validity with well-established measures. Age- and gender-specific population-based norms of the *ChEDE-Q8* allow for directly evaluating the severity of

children's eating disorder symptoms. Because children and adolescents between 7 and 18 years of age were analyzed, the *ChEDE-Q8* was proved to qualify for use over a wide age range. There is an ongoing debate about difficulties around the use of a questionnaire version of the *EDE* with younger individuals. For example, Goldschmidt and colleagues[23] argue that younger children may not be able to understand complex items *of ChEDE-Q* without assistance. We agree that in younger children the interview-based assessment of eating disorder psychopathology is preferable. On the other hand, face-to-face interviews cannot be applied in all areas (e.g., epidemiological studies or longitudinal studies with many assessment points) due to a lack of time or for financial reasons. There are thus some arguments in favor of conducting the *ChEDE-Q8* as a self-report assessment of eating disorder psychopathology in children. First, the questionnaire is relatively short and therefore appears suitable for young children. Second, the instrument allows for a reliable assessment in younger age groups (<12 years: α=.87). Third and most importantly, confirmation of strict measurement invariance indicated that the *ChEDE-Q8* measures the same construct across different children's age groups, which is a prerequisite for meaningful cross-group comparisons and emphasizes high measurement quality. However, in future studies, results obtained using the *ChEDE-Q8* should be compared to the results of a diagnostic interview (e.g., *ChEDE*).

Limitations

In order to achieve sufficient sample size, two different samples were combined for analysis. Although both samples were unselected, included boys and girls, and covered a broad age and socioeconomic range, measurement invariance analyses were conducted to account for these heterogeneities. Furthermore, it is unclear whether the psychometric properties of the *ChEDE-Q8* would also apply to child or adolescent samples with more severe eating disorder psychopathology or to samples from other countries. This is certainly an important area for future research, along with the need for examining predictive and

discriminant validity of the *ChEDE-Q8*. Because the *ChEDE-Q*8 was derived from the adult *EDE-Q*8,[12] there may be debate about whether age-specific features of eating disorder psychopathology are adequately represented. However, additional exploratory factor analyses produced an item selection identical with that found for the adult *EDE-Q8* (data not shown). Finally, although the *ChEDE-Q8* targets children ages 8-14 years, in accordance with the *ChEDE-Q*, the present study included somewhat younger children and older adolescents. The fact that the *ChEDE-Q8* showed favorable psychometrics for these broadened age ranges, however, highlighted its suitability and comparability across childhood and adolescence, so that methodological variances (e.g., based on changes in assessments) can be reduced. While the *ChEDE-Q8* targets at the assessment of eating disorder psychopathology related to AN, BN, and BED, just as the *ChEDE-Q*, alternative measures are needed to identify symptoms of ARFID, pica, or rumination disorder, e.g. the *Eating Disturbances in Youth-Questionnaire*.[42]

Conclusions

Although further validation studies are needed, particularly with respect to discriminant and predictive validity, the results suggested that the *ChEDE-Q8* has a clear unidimensional factor structure and favorable psychometrics, and thus represents an evidence-based, economical self-report assessment of eating disorder psychopathology in children. While the *ChEDE-Q8* appears to be particularly suitable within an epidemiological framework, the assessment of global eating disorder psychopathology, unifying restraint as well as eating, weight and shape concern, provides clinically relevant information on the level of eating disturbances in youth.

Acknowledgement

Table 1

Sociodemographic characteristics

Characteristics	Sample 1 (*n*=594)			Sample 2 (*n*=1242)			Total Sample (*N*=1836)		
Gender	*n* (%)			*n* (%)			*n* (%)		
Boys	290 (48.8)			637 (51.3)			927 (50.5)		
Girls	304 (51.2)			605 (48.7)			909 (49.5)		
Age	*M* (*SD*)	median	range	*M* (*SD*)	median	range	*M* (*SD*)	median	range
	10.81 (1.48)	11.00	8.00-16.81	11.95 (2.25)	11.80	7.12-17.99	11.58 (2.10)	11.30	7.12-17.99
Age group	*n* (%)			*n* (%)			*n* (%)		
7 to 10 y	117 (19.7)			265 (21.3)			382 (20.8)		
>10 to 12 y	292 (49.2)			387 (31.2)			676 (37.0)		
>12 to 14 y	179 (30.1)			348 (28.0)			527 (28.7)		
>14 to 18 y	6 (1.0)			242 (19.5)			248 (13.5)		

(*continued*)

Table 1 (continued)

Sociodemographic characteristics

Characteristics	Sample 1	Sample 2	Total Sample
	(*n*=594)	(*n*=1242)	(*N*=1836)
Weight status	*n* (%)	*n* (%)	*n* (%)
Underweight (BMI-SDS≤-1.28)		97 (7.8)	97 (5.3)
Normal weight (-1.28>BMI-SDS<1.28)		791 (63.7)	791 (43.1)
Overweight (1.28≤BMI-SDS<1.88)		107 (8.6)	107 (5.8)
Obesity (BMI-SDS≥1.88)		218 (17.6)	218 (11.9)
Missing	594 (100)	29 (2.3)	623 (34.0)

Note. Sample 1 = University of Marburg; Sample 2 = University of Leipzig; BMI-SDS = body mass index-standard deviation score. BMI-SDS is a measure of relative weight adjusted for child's age and sex.[35] BMI-SD scores for Sample 2 were determined based on objectively measured weight and height. As for Sample 1 only subjective weight and height were available, BMI-SD scores were not calculated.

Table 2

Means (M), standard deviation (SD), item difficulty (p_i), selectivity (r_{it}), and gender differences for the ChEDE-Q8 items and global score

	Total Sample				Boys				Girls				Group Differences			
	M	*SD*	p_i	r_{it}	*M*	*SD*	p_i	r_{it}	*M*	*SD*	p_i	r_{it}	*t*	*df*	*p*	d
Restraint over Eating	0.97	1.67	16	.57	0.91	1.65	15	0.55	1.04	1.69	17	.59	-1.73	1834	.085	0.04
Food Avoidance	0.63	1.33	11	.55	0.58	1.31	10	0.52	0.68	1.36	11	.57	-1.53	1834	.126	0.04
Preoccupation with Food	0.56	1.34	9	.43	0.59	1.40	10	0.43	0.53	1.27	9	.49	0.85	1834	.394	0.02
Feelings of Fatness	1.46	2.15	24	.80	1.24	2.02	21	0.77	1.69	2.26	28	.82	-4.45	1834	<.001	0.10
Desire to Lose Weight	1.59	2.33	27	.79	1.41	2.25	24	0.75	1.78	2.40	30	.82	-3.40	1834	.001	0.08
Guilt about Eating	0.68	1.32	11	.65	0.54	1.17	9	0.62	0.83	1.44	14	.67	-4.71	1834	<.001	0.11
Dissatisfaction with Weight	1.43	1.99	24	.78	1.19	1.89	20	0.75	1.68	2.07	28	.80	-5.28	1834	<.001	0.12
Discomfort Seeing Body	1.06	1.74	18	.75	0.86	1.58	14	0.71	1.25	1.87	21	.78	-4.86	1834	<.001	0.11
Global score	1.05	1.33	18	-	0.92	1.23	15	-	1.19	1.41	20	-	-4.37	1834	<.001	0.10

Table 3

Percentiles for the ChEDE-Q8 by age and gender

ChEDE-Q8 global score	Total Sample 7-18y (*n*=1836)	Boys					Girls				
		7-18y (*n*=927)	7-10y (*n*=185)	>10-12y (*n*=366)	>12-14y (*n*=265)	>14-18y (*n*=111)	7-18y (*n*=909)	7-10y (*n*=197)	>10-12y (*n*=313)	>12-14y (*n*=262)	>14-18y (*n*=137)
0.00	30	33	33	32	31	40	27	29	33	24	18
0.13	39	43	42	42	42	49	35	41	43	29	22
0.25	45	49	49	48	49	55	41	48	49	36	26
0.38	50	54	54	53	53	60	47	55	54	40	29
0.50	54	58	58	56	56	64	51	59	58	45	33
0.63	58	61	60	61	59	67	54	62	63	49	36
0.75	61	65	66	65	62	67	57	65	66	50	39
0.88	64	68	70	69	64	69	60	67	69	53	43
1.00	66	69	71	71	66	71	62	70	70	57	45
1.13	68	72	74	72	69	75	65	73	71	58	50
1.25	71	74	77	75	71	76	67	75	74	58	53
1.38	72	76	80	76	73	77	67	76	74	59	54
1.50	74	78	81	78	75	77	70	80	75	61	57
1.63	75	79	82	81	76	78	71	83	76	62	58
1.75	77	81	83	82	79	78	72	84	78	62	61
1.88	78	82	84	83	80	80	74	85	79	64	61
2.00	79	83	85	84	80	81	75	86	80	66	65
2.13	80	84	85	85	81	82	76	87	80	68	66
2.25	82	86	88	87	83	82	78	89	82	71	69
2.38	83	87	**90**	88	85	82	79	**90**	83	71	70
2.50	84	88	**91**	88	87	83	81	**90**	85	74	72
2.63	85	89	**91**	89	88	85	82	**91**	86	74	74
2.75	86	89	**91**	90	88	86	83	**91**	87	76	76
2.88	87	**90**	**91**	90	88	88	85	**92**	88	79	78
3.00	88	**91**	**92**	91	89	89	86	**93**	90	79	80
3.13	89	**91**	**92**	93	89	**90**	87	**93**	91	82	81
3.25	**91**	92	93	94	90	91	89	**95**	93	82	83
3.38	**92**	94	94	**95**	93	92	**90**	**95**	94	84	85
3.50	**93**	94	95	**95**	94	92	**91**	**95**	94	86	85
3.63	**94**	95	95	**95**	95	93	**92**	**97**	96	88	87
3.75	**95**	96	96	96	96	95	**93**	**97**	96	90	88
3.88	**95**	96	97	97	97	95	**94**	**97**	96	92	90
4.00	**96**	97	97	97	97	96	**95**	98	96	94	90
4.13	**97**	98	99	98	97	96	**96**	98	97	95	90
4.25	**97**	98	>99	98	97	97	**97**	99	97	97	92
4.38	**98**	98	>99	98	98	97	**97**	99	97	97	92
4.50	**98**	99	>99	98	99	98	**97**	99	98	97	93
4.63	**98**	99	>99	99	99	98	98	99	98	98	95
4.75	**99**	99	>99	99	99	98	98	99	99	99	96
4.88	**99**	99	>99	99	99	99	99	>99	99	99	96
5.00	**99**	99	>99	99	99	99	99	>99	>99	99	96
5.13	**99**	99	>99	99	99	99	99	>99	>99	99	96
5.25	**99**	>99	>99	>99	99	>99	99	>99	>99	99	**97**
5.38	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	99
5.50	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	99
5.63	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	>99
5.75	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	>99
5.88	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	>99
6.00	**>99**	>99	>99	>99	>99	>99	>99	>99	>99	>99	>99

Note. Normative data are presented as *ChEDE-Q8* global mean scores with corresponding percentiles. Percentiles are shown for the total sample and for subsamples based on age and gender; light grey area indicate critical severity of eating disorder psychopathology; grey area indicate high severity of eating disorder psychopathology; dark grey area indicate very high severity of eating disorder psychopathology.

Appendix A

Analysis of factorial invariance using multigroup confirmatory factor analyses

	χ^2	df	CFI	ΔCFI	RMSEA	ΔRMSEA	Measurement Invariance Test[a]
Gender (boys; girls)							
Configural invariance	531.49	36	.941	-	.079	-	
Weak invariance (equal loadings)	544.12	43	.940	−.001	.073	−.006	
Strong invariance (equal loadings + intercepts)	564.48	46	.936	−.04	.073	.000	✓
Strict invariance (equal loadings + intercepts + residuals)	606.06	54	.938	+.002	.066	-.007	✓
Age (7-10y; >10-12y; >12-14y; >14-18y)							
Configural invariance	555.31	72	.947	-	.075	-	
Weak invariance (equal loadings)	611.65	93	.943	−.004	.069	−.006	✓
Strong invariance (equal loadings + intercepts)	661.23	102	.935	−.008	.070	+.001	✓
Strict invariance (equal loadings + intercepts + residuals)	848.36	126	.929	−.006	.066	−.004	✓

(*continued*)

Appendix A (continued)

*Analysis of factorial invariance using multigroup confirmatory factor analyses*

	χ^2	*df*	CFI	ΔCFI	RMSEA	ΔRMSEA	Measurement Invariance Test[a]
Weight status (overweight; normal weight)							
Configural invariance	401.60	36	.924	-	.090	-	
Weak invariance (equal loadings)	417.72	43	.916	-.008	.087	-.003	✓
Strong invariance (equal loadings + intercepts)	419.62	46	.916	-.000	.084	-.003	✓
Strict invariance (equal loadings + intercepts + residuals)	688.75	54	.901	-.015	.084	.000	✓
Subsamples (sample 1; sample 2)							
Configural invariance	506.97	36	.945	-	.078	-	✓
Weak invariance (equal loadings)	539.16	43	.941	−.004	.074	−.004	✓
Strong invariance (equal loadings + intercepts)	552.00	46	.938	−.003	.073	-.001	✓
Strict invariance (equal loadings + intercepts + residuals)	613.79	54	.937	−.001	.069	-.004	✓

Note. CFI=Comparative Fit Index; ΔCFI=differences between models (1 and 2; 2 and 3; 3 and 4); RMSEA=root mean square error of approximation; ΔRMSEA=differences between models (1 and 2; 2 and 3; 3 and 4); [a]*ΔCFI≤−.010* supplemented by *ΔRMSEA≥.015* indicates non-invariance. √ indicates invariance

Appendix B.

Confirmatory factor analysis of the higher order general factor model.
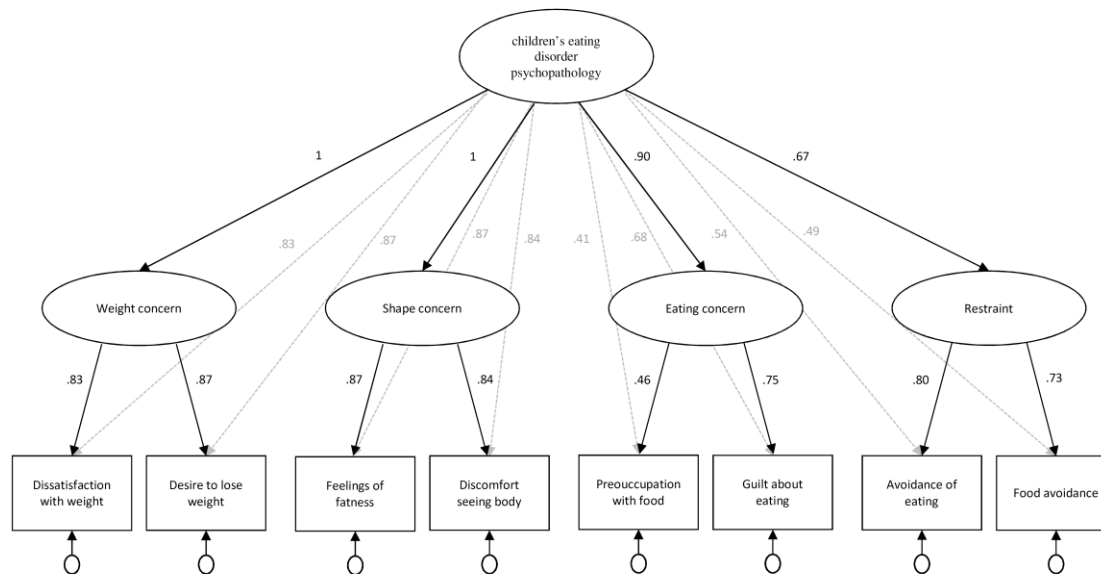
Note: The model was estimated using MLM estimation. Loadings are standardized. Grey loading indicates the second-order corrected factor loadings (item loading * one-order factor loading).

References

1. Hölling H, Schlack R. Essstörungen im Kindes-und Jugendalter: Erste Ergebnisse aus dem Kinder- und Jugendgesundheitssurvey (KiGGS). [Eating disorders in children and adolescents. First results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS)]. *Bundesgesundhbl* 2007;50:794–799.

2. Tanofsky-Kraff M, Yanovski SZ, Wilfley DE, Marmarosh C, Morgan CM, Yanovski JA. Eating-disordered behaviors, body fat, and psychopathology in overweight and normal-weight children. J *Consult Clin Psychol* 2004;*72*:53–61.

3. Herpertz-Dahlmann B, Wille N, Hölling H, Vloet TD, Ravens-Sieberer U. Disordered eating behaviour and attitudes, associated psychopathology and health-related quality of life: results of the BELLA study. *Eur Child Adolesc Psy* 2008;17:82–91.

4. Boutelle K, Neumark-Sztainer D, Story M, Resnick M. Weight control behaviors among obese, overweight, and nonoverweight adolescents. *J Pediatr Psychol* 2002;27:531–540.

5. Jendrzyca A, Warschburger P. Weight stigma and eating behaviours in elementary school children: A prospective population-based study. *Appetite* 2016;102:51–9.

6. Kurz S, van Dyck Z, Dremmel D, Munsch S, Hilbert A. Early-onset restrictive eating disturbances in primary school boys and girls. *Eur Child Adoles Psy* 2015;24:779–785.

7. Cooper PJ, Watkins B, Bryant-Waugh R, Lask B. The nosological status of early onset anorexia nervosa. *Psychol Med* 2002;32:873-880.

8. Treasure J, Claudino AM, Zucker N. Eating disorders. *Lancet* 2010;375:583–593.

9 Cardona Cano S, Tiemeier H, van Hoeken D, Tharner A, Jaddoe VWV, Hofman A et al. Trajectories of picky eating during childhood: A general population study. *Int J Eat Disord* 2015;48:570–579.

10. Kotler LA, Cohen P, Davies M, Pine DS, Walsh TB. Longitudinal relationships between childhood, adolescent, and adult eating disorders. *J Am Acad Child Adolesc Psychiatry* 2001;40:1434–1440.

11. Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US Children in the 2001-2004 NHANES. *Pediatr* 2010;125:75–81.

12. Field AE, Sonneville KR, Crosby RD, Swanson SA, Eddy KT, Camargo CA et al. Prospective Associations of Concerns About Physique and the Development of Obesity, Binge drinking, and drug use among adolescent boys and young adult men. *JAMA Pediatr* 2014;168:34.

13. Sonneville KR, Calzo JP, Horton NJ, Haines J, Austin SB, Field AE. Body satisfaction, weight gain and binge eating among overweight adolescent girls. *Int J Obes Relat Metab Disord* 2012;36:944–949.

14. Hilbert A, Brauhardt A. Childhood loss of control eating over five-year follow-up. *Int J Eat Disord* 2014;47:758–761.

15. Sonneville KR, Horton NJ, Micali N, Crosby RD, Swanson SA, Solmi F et al. Longitudinal associations between binge eating and overeating and adverse outcomes among adolescents and young adults. *JAMA Pediatr* 2013;167:149.

16. Field AE, Sonneville KR, Micali N, Crosby RD, Swanson SA, Laird NM et al. Prospective association of common eating disorders and adverse outcomes. *Pediatr* 2012;130:e289-e295.

17. Micali N, Solmi F, Horton NJ, Crosby RD, Eddy KT, Calzo JP et al. Adolescent eating disorders predict psychiatric, high-risk behaviors and weight outcomes in young adulthood. *J Am Acad Child Adolesc Psychiatry* 2015;54:652-659.

18. Herpertz-Dahlmann B, Dempfle A, Konrad K, Klasen F, Ravens-Sieberer U. Eating disorder symptoms do not just disappear: the implications of adolescent eating-disordered behaviour for body weight and mental health in young adulthood. *Eur Child Adolesc Psy* 2015;24:675–684.

19. Herpertz-Dahlmann B, Wille N, Hölling H, Vloet TD, Ravens-Sieberer U. Disordered eating behaviour and attitudes, associated psychopathology and health-related quality of life: Results of the BELLA study. *Eur Child Adolesc Psychiatry* 2008;17:82–91.

20. Fairburn CG, Beglin SJ. Assessment of eating disorders: Interview or self-report questionnaire? *Int J Eat Disord* 1994;16:363–370.

21- Decaluwé V. Child Eating Disorder Examination-Questionnaire. Dutch translation and adaptation of the Eating Disorder Examination-Questionnaire, authored by CG Fairburn and SJ Beglin, unpublished manuscript 1999.

22. Decaluwé V, Braet C. Assessment of eating disorder psychopathology in obese children and adolescents: Interview versus self-report questionnaire. *Beha Res Ther* 2004;42:799–811.

23. Goldschmidt AB, Doyle AC, Wilfley DE. Assessment of binge eating in overweight youth using a questionnaire version of the Child Eating Disorder Examination with instructions. *Int J Eat Disord* 2007;40:460–467.

24. Hilbert A, Hartmann AS, Czaja J. Child Eating Disorder Examination-Questionnaire: Evaluation der deutschsprachigen Version des Essstörungsfragebogens für Kinder Klinische Diagnostik und Evaluation [Evaluation of the german-speaking version of the Child Eating Disorder Examination Questionnaire]. *Klin Diag Eval* 2008:447–463.

25. Kliem S, Mößle T, Zenger M, Strauß B, Brähler E, Hilbert A. The Eating Disorder Examination-Questionnaire 8: A brief measure of eating disorder psychopathology (EDE-Q8). *Int J Eat Disord* 2015;49:613-616.

26. Berger U, Hentrich I, Wick K, Bormann B, Brix C, Sowa M, et al. Psychometric quality of the "Eating Attitudes Test"(German Version EAT-26D) for measuring disordered eating in pre-adolescents and proposal for a 13-Item Short Version. *Psychother Psych Med* 2012; 62: 223-226.

27. Hilbert A, Czaja J. Essanfälle und Adipositas im Kindesalter [Binge eating and childhood obesity]. *Psychother Psych Med* 2007;57:413–419.

28. Quante M, Hesse M, Döhnert M, Fuchs M, Hirsch C, Sergeyev E et al. The LIFE child study: A life course approach to disease and health. *BMC Public Health* 2012;12:e1-e14.

29. Mendelson BK, Mendelson MJ, White DR. Body-Esteem Scale for Adolescents and Adults. *J Pers Assess* 2001;76:90–106.

30. Goodman R. The Strengths and Difficulties Questionnaire: A research note. *J Child Psychol Psychiat* 1997;38:581–586.

31. Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V et al. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. [Percentiles of body mass index in children and adolescents evaluated from different regional German studies]. *Monatsschr Kinderheilkd* 2001;149:807–818.

32. Hilbert A, Buerger A, Hartmann AS, Spenner K, Czaja J, Warschburger P. Psychometric evaluation of the Eating Disorder Examination adapted for children. *Eur Eat Disorders Rev* 2013;21(4):330–339.

33. Field A. Discovering statistics using SPSS. 3rd ed. Los Angeles, Calif.: Sage; 2011

34. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–399.

35. Chou CP, Bentler PM, Satorra A. Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: a Monte Carlo study. *Br J Math Stat Psychol* 1991;44:347–357.

36. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.

37. Meredith W, Teresi JA. An essay on measurement and factorial invariance. *Med Care* 2006;44:69-77.

38. Millsap RE, Kwok O. Evaluating the impact of partial factorial invariance on selection in two populations. *Psych Meth* 2004;9(1):93–115.

39. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Modeling* 2007;14:464–504.

40. Rosseel Y. Lavaan: An R package for structural equation modeling. *J Stat Softw* 2012;48(2):1-36.

41. van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1-67.

42. Kurz S, Van Dyck Z, Dremmel D, Munsch S, Hilbert A. (2015). Early-onset restrictive eating disturbances in primary school boys and girls. *Eur Child Adolesc Psychiatry*, 24: 779-785.