Universität Potsdam

Humanwissenschaftliche Fakultät

Jana Zitzmann | Petra Warschburger

Psychometric Properties of Figure Rating Scales in Children

The Impact of Figure Ordering

Suggested citation referring to the original publication:
Journal of Personality Assessment (2018)
DOI http://dx.doi.org/10.1080/00223891.2018.1505731
ISSN (print) 0022-3891
ISSN (online) 1532-7752

JOURNAL OF PERSONALITY ASSESSMENT
https://doi.org/10.1080/00223891.2018.1505731

Psychometric Properties of Figure Rating Scales in Children: The Impact of Figure Ordering

Jana Zitzmann and Petra Warschburger

Department of Psychology, University of Potsdam, Potsdam, Germany

ABSTRACT

This study examined psychometric properties of figure rating scales, particularly the effects of ascending silhouette ordering, in 153 children, 9 to 13 years old. Two versions of Collins's (1991) figural rating scale were presented: the original scale (figures arranged ascendingly) and a modified version (randomized figure ordering). Ratings of current and ideal figure were elicited and body dissatisfaction was calculated. All children were randomly assigned to one of two subgroups and completed both scale versions in a different sequence. There were no significant differences in figure selection and body dissatisfaction between the two figure orderings. Regarding the selection of the current figure, results showed that girls are more affected by the silhouette ordering than boys. Our results suggest that figure rating scales are both valid and reliable, whereby correlation coefficients reveal greater stability for ideal figure selections and body dissatisfaction ratings when using the scale with ascending figure ordering.

ARTICLE HISTORY
Received 20 December 2017
Revised 6 July 2018

In our society today, media presentations of thin models and appeals to diet are omnipresent and thus easily accessible even to young children. These messages can influence our mental representations of our own bodies (i.e., our body image). Body image disturbance is described as a multidimensional concept, including perceptual, affective, cognitive, and behavioral aspects (Thompson, 1995; Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999). As a component of body image disturbance, body dissatisfaction encompasses negative attitudes and evaluations regarding one's own body (Garner & Garfinkel, 1981). Recent studies have observed that body dissatisfaction is already highly prevalent among young children (e.g., Dion et al., 2016; Tatangelo, McCabe, Mellor, & Mealey, 2016; Tremblay, Lovsin, Zecevic, & Lariviere, 2011). Taking into account that body dissatisfaction plays an important role in the development and maintenance of eating and weight-related disorders (e.g., Paxton & Damiano, 2017; Thompson, 1995; Thompson et al., 1999), childhood could be a sensitive period for prevention and treatment approaches (Dion et al., 2016; Tatangelo et al., 2016; Thompson et al., 1999). To develop such approaches, reliable and valid assessment tools suitable for children are needed.

Numerous assessment tools are available to evaluate body dissatisfaction. Most commonly, figural drawing scales are applied to measure body dissatisfaction (e.g., Furnham & Alibhai, 1983; Gardner & Brown, 2010a, 2011; Thompson & Gray, 1995; Truby & Paxton, 2002). From a series of frontal images, usually ranging from lower to higher body weight, individuals are asked to select the figure that best represents their current and ideal body size. Body dissatisfaction is then operationalized as the discrepancy between these two ratings. Figural drawing scales provide several advantages. Besides their high face validity, they can be applied in group settings and answered easily and quickly. In comparison to questionnaires, fewer verbal abilities are required, which makes their application particularly suitable for studies involving younger children (Gardner & Brown, 2011; Thompson & Altabe, 1991; Truby & Paxton, 2002). Furthermore, it has been shown that figure ratings and body dissatisfaction go along with weight changes (e.g., Sala et al., 2012; Song et al., 2006; Walker, Gately, Bewick, & Hill, 2003), supporting the ecological validity of these measures. In pediatric research, the figural drawing scale developed by Collins (1991) for the application with preadolescent children is probably the most commonly used (Smolak, 2004).

Although many sets of figural drawing scales for children have been developed, for the majority of these scales, little or no psychometric data have been documented (Smolak, 2004; see Gardner & Brown, 2010a, for a review). Furthermore, several possible methodological shortcomings have been mentioned (see Gardner & Brown, 2010a, for a review). Probably the most important point concerns the arrangement of figures within the scale. In many cases, all figural drawings are presented on a single sheet of paper and the figures are arranged in an ascending sequence of

CONTACT Petra Warschburger warschb@uni-potsdam.de Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.

sizes from left (thin) to right (obese; Gardner & Brown, 2010a; e.g., Collins, 1991; Fallon & Rozin, 1985; Rand & Wright, 2001; Stunkard, Sorensen, & Schulsinger, 1983). This arrangement is being discussed to cause a reporting bias. Regarding a Likert-type scale with ascending order, Nicholls, Orr, Okubo, and Loftus (2006) found a bias to respond to the left of the scale's midpoint. A similar effect was stated by Gardner (1996), summarizing results of studies in which a video distortion technique was used to examine the perceptual aspect of body image. Judgments of one's body size were smaller when participants were confronted with an obviously undersized presentation of the body proportions compared to an obviously oversized presentation (e.g., Gardner, Espinoza, Urrutia, Morrell, & Gallegos, 1989; Gardner, Martínez, & Espinoza, 1987; Gardner, Martínez, Espinoza, & Gallegos, 1988). Since figural drawing scales typically use an ascending sequence, a bias toward thinner figures might be expected (Gardner & Brown, 2010b, 2011).

However, research explicitly comparing different figure orderings is scarce. Previous research on the influence of figure ordering among adolescents or adults (e.g., Doll, Ball, & Willows, 2004; Duncan, Dodd, & Al-Nakeeb, 2005; Gardner & Brown, 2011; Paludo et al., 2011) yielded inconsistent findings. Gardner and Brown (2011) demonstrated that bigger figures were selected to represent the participants' perceived body and less body dissatisfaction was reported when using a scale with images in ascending order compared to a scale with randomly ordered images. In contrast, Paludo et al. (2011) reported evidence for the reverse direction. In their study, participants expressed being more dissatisfied with their bodies when judging figures presented in ascending order. Other studies found no influence of the presentational method on current and ideal body size (Doll et al., 2004; Duncan et al., 2005) or body dissatisfaction (Duncan et al., 2005). To sum up, the few studies addressing the influence of figure ordering within figural drawing scales have led to inconsistent findings. All studies concentrated on scales developed for adolescent or adult participants and can therefore not be generalized to children. To the best of our knowledge, the influence of figure ordering among children has not yet been analyzed.

In addition, there is little research on the criterion validity of figural drawing scales, especially for their use in childhood and depending on the figure ordering. With respect to ascending figure presentation, previous studies have underscored the criterion validity of figural drawing scales for their use in children. For example, Collins (1991) found a significant relationship between children's (on average 8 years old) body size perception using the figure rating scale and their body mass index (BMI; $r = .37$). Similar results ($r = .54$) were reported by Coelho, Padez, Moreira, Rosado, and Mourão-Carvalhal (2013) in a group of 7- to 10-year-old children. Age and gender should be considered as moderating influences as well, but there exist only limited data on that topic. Lombardo, Battagliese, Pezzuti, and Lucidi (2014) reported a significant relationship between the selected current figure and the age-corrected BMI, with higher correlations among older children (third grade or older, $r \geq .45$) compared to younger children (second grade or younger, $r \leq .31$). Furthermore, Truby and Paxton (2002) demonstrated higher correlations among girls ($r = .60$, 10–12 years old) compared to boys ($r = .35$, 10–12 years old). However, Williamson and Delin (2001) found no age- or gender-related differences in the accuracy of self-perception in 5- to 10-year-old children. With respect to the criterion validity of figural drawing scales with randomized figure ordering, no studies have been carried out with children. However, initial evidence comes from studies with adult participants: Significant correlations between figure selection and BMI were observed when silhouette drawings were arranged sequentially ($r = .71$, Gardner & Brown, 2011) and randomly ($r = .67$, Gardner & Brown, 2010b; $r = .76$, Gardner, Jappe, & Gardner, 2009). Although criterion validities of both scale versions were comparable, a significantly more pronounced overestimation of one's body size (+4.8%) was observed when placing figures sequentially (Gardner & Brown, 2011).

Further, there is consistent evidence for the construct validity of figure rating scales in children. Truby and Paxton (2002) reported significant relationships between the current–ideal discrepancy on the Children's Body Image Scale (CBIS) and the items "I think I am too thin/fat" (girls, $r = .47$; boys, $r = .51$), "I would like to be thinner/fatter" (girls, $r = -.64$; boys, $r = -.56$), in 8- to 12-year-old children. In addition, Wertheim, Paxton, and Tilgner (2004) found significant correlations between body dissatisfaction assessed by the Contour Drawing Rating Scale (CDRS) in adolescent girls (on average 13.21 years old), indicating construct validity. Sherman, Iacono, and Donnelly (1995) reported high correlations between two different figural drawing scales in 11-year-old ($r = .81$) and 17-year-old ($r = .85$) participants. These studies support the construct validity of the figural rating scales. However, all of them used figure presentations in an ascending order. There is no data on other presentation formats.

Besides the validity of the figure drawing scales, another criticism has been that the ascending order might lead to spuriously high test–retest reliability, as subjects have less difficulty remembering the previously marked figures (e.g., Gardner, Friedman, & Jackson, 1998). On the one hand, Wertheim et al. (2004) reported strong test–retest reliabilities of figure selections representing current ($r = .84$) and ideal ($r = .78$) body size as well as of body dissatisfaction ($r = .82$) over 2 weeks, regarding the CDRS with ascending figure ordering among early adolescent girls (on average 13.21 years old). On the other hand, there are initial results reporting comparable test–retest reliabilities over 3 weeks ($r \geq .87$) for randomized silhouette ordering among adults (Gardner, Stark, Jackson, & Friedman, 1999). However, until now no study has directly compared test–retest reliabilities when scales with randomized and ordered figure presentation are applied.

The availability of reliable and sound assessment tools for body dissatisfaction is a mandatory requirement in prevention and intervention research. Despite the fact that figure ratings are most commonly used in research, there are only limited data on their psychometric properties, especially among children. Particularly the question of whether the ascending order of figures is associated with biased estimations has not yet been studied. Therefore, this study aimed to investigate the psychometric properties of the probably most widely used figural drawing scale in research with children (Smolak, 2004). First, we wanted to examine whether figure selections (current, ideal) and body dissatisfaction differ as a function of figure presentation (in random vs. ascending order). Second, by comparing current figure selection and BMI as well as body dissatisfaction assessed by the figural drawing scales and by the corresponding Eating Disorder Inventory for Children (EDI-C) subscale, criterion and convergent validity depending on the figure ordering could be measured. Third, test–retest reliabilities of the different presentation modes were explored. It was hypothesized that (a) selections of current and ideal figure and the resulting body dissatisfaction, (b) criterion and convergent validity coefficients, and (c) test–retest reliability coefficients are comparable for both presentational methods. In addition, age-, gender- and weight-related differences were explored.

Table 1. Demographic characteristics of the two subsamples and group comparison.

Characteristics	Subsample 1[a]	Subsample 2[b]	Test statistics
Age			$t(151) = 0.685$, $p = .494$
M	10.51	10.39	
SD	1.00	1.04	
Range	9–13	9–13	
Gender			$\chi^2(1) = 0.003$, $p = .958$
Male			
n	53	23	
%	49.5	50.00	
Female			
n	54	23	
%	50.5	50.00	
BMI			$t(60.73) = -1.227$, $p = .225$
M	18.39	19.30[c]	
SD	3.35	4.38[c]	
Range	13.26–29.50	12.78–33.02[c]	
Weight category			$\chi^2(2) = 5.717$, $p = .057$
Underweight			
n	11	7	
%	10.28	16.67[c]	
Normal weight			
n	82	24	
%	76.64	57.14[c]	
Overweight			
n	14	11	
%	13.08	26.19[c]	

Note. BMI = body mass index.
[a]*n* = 107. [b]*n* = 46. [c]For 4 children, no anthropometric data are available.

Method

Participants

Participants were recruited in two waves of data collection in 2013 and 2014. The first subsample consisted of 107 preadolescent children, recruited from two elementary schools in Berlin, and included the assessment of test–retest reliability. The second data collection included another 46 school-aged children, recruited from another elementary school in Berlin. Hence, the final sample consisted of 153 children aged 9 to 13 years ($M = 10.48$, $SD = 1.01$) attending grade levels 4 through 6. There was a high response rate (70.06%). Reasons for declined participation were missing informed consent from the parents or the children as well as absence from school (e.g., due to illness). The reason for dropout from the first to the second time of measurement was absence from school (e.g., due to illness). The gender distribution was evenly balanced (50.33% girls) and the majority of the children (71.14%) were normal weight. No significant group differences between the two assessment points were observed. Demographic characteristics including gender, age, BMI, and weight categories for both subsamples as well as the results of the group comparisons are shown in Table 1.

Measures

Demographics and BMI

Children reported on their gender and age. Height and weight were assessed by the investigators with standardized equipment. Height was measured using ADE Ultrasound Height Measuring Unit MZ10020, to the nearest 0.01 m, and weight was measured using Waagen-Schrenk Model MS-4202L, to the nearest 0.01 kg, both without shoes. On that basis, BMI (kg/m^2) scores were calculated. To account for the maturing process, age- and gender-adjusted BMI standard deviation scores (BMI–SDS) were calculated according to German reference data (Kromeyer-Hauschild et al., 2001). The BMI–SDS indicates the difference from the median BMI in standard deviation units. In accordance with those national guidelines recommended by Kromeyer-Hauschild et al. (2001), weight categories were defined as underweight (BMI <10th percentile), normal weight (10th percentile ≤ BMI ≤90th percentile), and overweight (BMI >90th percentile).

Figural Rating Scale

The Figural Rating Scale developed by Collins (1991) was used to assess children's perceived current and desired ideal body size, as well as to measure their degree of body dissatisfaction. The gender-specific scale comprises seven male and female child figure drawings illustrating different body weights. Figures are presented in an ascending order ranging from the leanest silhouette on the left side (1) to the heaviest silhouette on the right side (7). In the first step, children were asked to select the figure that best represented their own body size (current), and in the second step the figure that represented how they would like to look (ideal). Body dissatisfaction was then conceptualized as a discrepancy between the two responses (current–ideal). Hence, higher values represent more pronounced body dissatisfaction. The positive and negative signs indicate the direction of the body dissatisfaction—the desire to become larger or thinner. For the statistical analyses, absolute (unsigned) values of body dissatisfaction were used. Collins (1991) reported a 3-day test–retest reliability of $r = .71$ for actual and $r = .59$ for ideal body size. Furthermore, criterion validity was supported by a significant positive correlation of the selected figure with the measured BMI ($r = .37$; Collins, 1991).

Modified figural rating scale

Based on the original scale (Collins, 1991), a modified version was generated in which the seven male and female child figures were arranged randomly. Due to separate randomizations, the order of the figures differs slightly between the two subsamples. As previous analyses have revealed, no significant differences between the two randomized scale versions were found and the data were merged.

EDI–C: Body Dissatisfaction subscale

For convergent validation of the current–ideal discrepancy, the Body Dissatisfaction subscale of the German version of the EDI for children (EDI–C; Thiels, Salbach-Andrae, Bender, & Garner, 2011; original, Garner, 1991) was applied. This subscale aims to assess an individual's dissatisfaction with different body parts. Participants were asked how satisfied they were regarding the overall size of their body, as well as with particular parts of their body. The scale consists of 9 items (e.g., "I am satisfied with my body") that are rated on a 6-point Likert scale ranging from 1 (*never*) to 6 (*always*). A sum score is computed to indicate the overall body dissatisfaction. To accommodate missing values, a mean score was also computed. The scale has good internal consistency, with $\alpha = .86$ (Thurfjell, Edlund, Arinell, Hagglof, & Engstrom, 2003). Furthermore, moderate correlations with the BMI ($r = .40$) give a first indication of convergent validity (Thurfjell et al., 2003). In this study, Cronbach's alpha for this scale was .81.

Procedures

The study design is depicted in Figure 1. All children were randomly assigned to one of two subgroups. In each subgroup, the two different methods of figure ordering (ascending, random) were tested, whereby the sequence in which the two scale versions were applied varied. While the first group initially retained the original scale, the second group first received the modified version. Children received a single sheet with the respective silhouettes and were asked to indicate their actual body size (current) and their ideal body size (ideal) on the presented paper. After completion of this task, the investigator collected the material. To minimize the risk of the children remembering what they had filled in, a 10-min break followed this test session. This distraction phase consisted of two different tasks. First, the children had to perform a section from the concentration task for third and fourth grades (German: Konzentrationstest für 3. und 4. Klassen, KT 3-4; Heck-Möhling, Reinhard, & Boehle, 1993), which was announced as a puzzle. On the top of the sheet, different views of dice were presented and the children were instructed to find the corresponding dice below the instructions. Second, a child-friendly relaxation exercise with a guided visualization was applied, whereby the children were taken on an imaginary journey with themes not related to their body. Afterward, the children received the second version of the figural drawing scale and the EDI–C. Finally, height and weight were measured anonymously by the investigators. To collect the retest data, there was a second appointment 3 days later within the first subsample. In this second session, every subgroup received the scale they had to answer first in the previous session (see Figure 1).

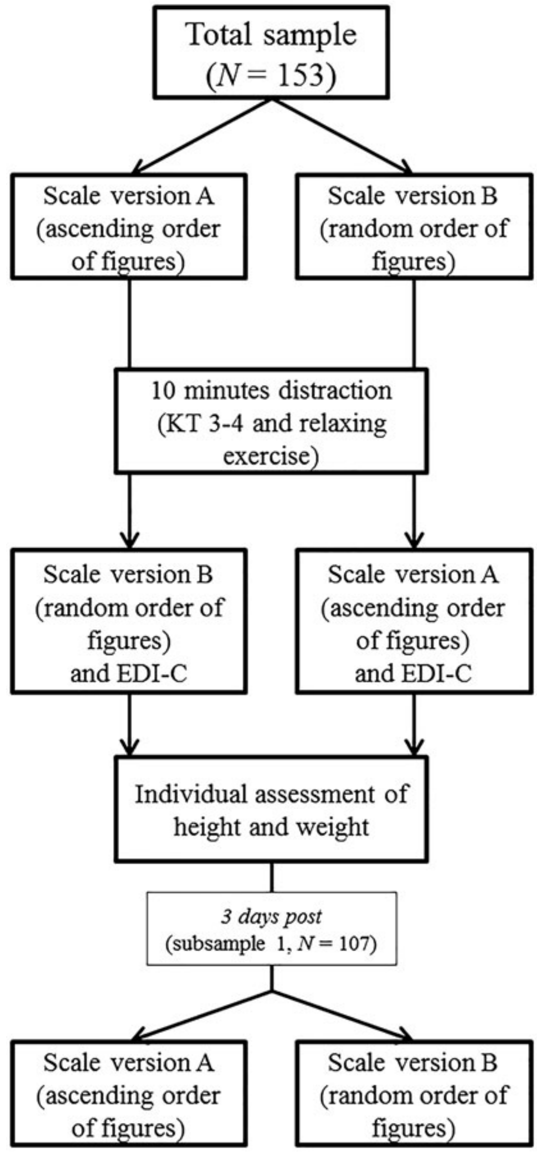

Figure 1. Study design.

In both sessions, data were collected inside the classroom. The children were provided with written instructions explaining the tasks and materials. For their participation, the children received small gifts (sweets). All participants volunteered for the study and parents gave their informed consent. The study was approved by the local ethics committee.

Statistical analysis

As there were no group differences with regard to demographic characteristics (see Table 1), the participants of both subsamples were merged. First, to determine the correspondence between ratings on both scale versions, intraclass correlations (ICCs) between figure selections and between body dissatisfaction values derived from the two scale versions were determined. The applied ICC design was a two-way random effects model (Shrout & Fleiss, 1979), not assuming interaction effects between the raters.

To compare the applied scale versions in terms of both figure selections (the children's current body size and their ideal body size) as dependent variables, a one-way multivariate analysis of variance (MANOVA, using Pillai's trace) was conducted to avoid inflating the risk of Type I error. In the case of significant effects, separate univariate analyses of variance (ANOVAs) were performed for each dependent variable. With respect to the reported body dissatisfaction, the ascending scale and the randomized scale were compared by conducting a one-way ANOVA. In a second step, influences of gender, age, and weight were investigated in an exploratory manner. For age and weight, two one-way multivariate analyses of covariance (MANCOVAs, using Pillai's trace) and two one-way analyses of covariance (ANCOVAs) were performed, considering age and the BMI–SDS as covariates in addition to figure ordering. For gender, a two-way MANOVA (using Pillai's trace) and a two-way ANOVA were performed, considering gender in addition to figure ordering. Additionally, the same analyses were conducted for weight categories (underweight, normal weight, overweight) to facilitate the interpretability of the results concerning weight. Due to our research questions, interaction effects were of special interest. In the case of significant main effects and interaction effects, post hoc tests or simple effects analyses with Bonferroni correction were performed. Effect sizes derived from focused comparisons were computed.

Criterion validity was determined by means of the correlation between current figure selection and BMI, depending on figure ordering. Convergent validity was measured by the correlation between body dissatisfaction assessed by the figural drawing scale (absolute value) and the corresponding EDI–C subscale, depending on figure ordering. Differences in criterion and convergent validity depending on the presentation method were tested using a t-statistic (see Chen & Popovich, 2002). The resulting value was checked against the critical value 1.98 ($p < .05$) with $N - 3$ degrees of freedom (df).

Finally, test–retest reliability was determined. For each scale version, the figure selections and values of body dissatisfaction obtained in Test Session 1 were correlated with those in Test Session 2. To examine differences in test–retest reliability due to figure ordering, a z score of the differences between these correlations was calculated (Field, 2015). Again, the z score and the corresponding p value were looked up in the table for the normal distribution.

A statistical power analysis (G*Power 3.1) was performed. The sample size of this study was shown to be adequate to detect small effects with a power of .95. The p value was set at $< .05$ for all analyses. Assumptions of all analyses were tested previously. Data analyses were performed using SPSS for Windows (Version 22.0).

Results

Comparison of ascending versus random figure ordering

Table 2 summarizes the data on the correspondence between the ordered and randomized silhouette presentations regarding selections of current and ideal figure as well as body dissatisfaction values. High ICCs and a high percentage of identical figure ratings ($\geq$ 84.97%) were observed.

In Table 3, means and standard deviations for figure ratings and body dissatisfaction are presented separately for the total sample as well as the gender- and weight-related subgroups of the statistical analyses reported here. In the first step, we investigated whether there were significant differences in figure selection and body dissatisfaction depending on the method of presentation. One-way MANOVA revealed no significant effect of figure arrangement on the selection of current and ideal figures, $V = 0.02$, $F(2, 151) = 1.27$, $p = .283$. Furthermore, one-way ANOVA comparing body dissatisfaction values revealed no significant difference between ascendant and random figure ordering, $F(1, 152) = 0.01$, $p = .913$, $d = 0.219$. To explore the potential influences of gender, age, and weight, we conducted a series of additional exploratory analyses. Due to our research questions, interaction effects were of particular interest and are primarily reported here.

Gender

With respect to gender, a MANOVA revealed no significant main effects (all $p > .05$). A significant interaction effect was observed, $V = 0.04$, $F(2, 150) = 3.43$, $p = .035$, indicating that the figure arrangement had different effects on children's ratings depending on their gender. Separate univariate ANOVAs on the outcome variables revealed a significant interaction effect between figure ordering and gender for the selection of the current figure, $F(1, 151) = 6.31$, $p = .013$, but not for the selection of the ideal figure ($p > .05$). To break down this interaction, simple effects analyses were performed, whereby the results revealed a significant difference between both scale versions for female ($p = .012$, $d = 0.292$) but not for male children ($p > .05$, $d = 0.118$). Girls were shown to select larger current figures using the randomized scale version compared to the ordered scale version (see Table 3). Regarding body dissatisfaction, the ANOVA results revealed neither significant main

Table 2. Correspondence between the ascending and randomized presentation methods regarding figure selection (current and ideal figure) and body dissatisfaction.

		Difference values[a]					
		Correspondence		Positive		Negative	
	ICC	n	%	n	%	n	%
Current figure	.888*	135	88.24	7	4.57	11	7.18
Ideal figure	.746*	130	84.97	8	5.23	15	9.80
Body dissatisfaction	.750*	130	84.97	10	6.54	13	8.50

Note: ICC = intraclass correlation. Correspondence indicates the same figure selections and body dissatisfaction values for both scale versions.
[a]Difference values are computed by subtracting values resulting from the randomized scale version from values resulting from the ascending scale version. A positive difference indicates larger figure selections and body dissatisfaction values using the ascending scale version and vice versa.
*$p < .001$.

Table 3. Means and standard deviations for figure ratings and body dissatisfaction in the total sample and in the gender- and weight-related subgroups of the statistical analyses, dependent on the presentational method (ascending vs. random order).

	Current figure				Ideal figure				Body dissatisfaction			
	Ascending order		Random order		Ascending order		Random order		Ascending order		Random order	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total sample	4.03	0.99	4.07	0.94	3.57	0.73	3.61	0.75	−0.46	1.07	−0.45	1.02
Gender[a]												
Female	4.08	1.00	4.21	0.89	3.62	0.56	3.69	0.57	−0.45	1.03	−0.52	0.94
Male	3.97	0.99	3.92	0.96	3.51	0.87	3.54	0.90	−0.46	1.11	−0.38	1.10
Weight category[b]												
Underweight	2.89	0.68	2.89	0.68	3.61	0.78	3.56	0.71	0.72	0.67	0.67	0.69
Normal weight	3.97	0.72	4.03	0.70	3.58	0.74	3.58	0.75	−0.40	0.80	−0.44	0.84
Overweight	5.04	0.98	5.08	0.76	3.56	0.65	3.72	0.79	−1.48	1.16	−1.36	0.95

[a]$n = 77$ female and $n = 76$ male. [b]$n = 18$ underweight, $n = 106$ normal weight, and $n = 25$ overweight.

effects of figure arrangement and gender nor a significant interaction effect between both variables (all $p > .05$).

Age

With respect to age, the results of the MANCOVA revealed no significant main effects of figure ordering and age (all $p > .05$). In addition, no significant interaction effect between both variables was observed ($p > .05$). In the case of body dissatisfaction, an ANCOVA revealed similar results. Neither the main effects nor the interaction effect were shown to be significant (all $p > .05$).

BMI–SDS

Using MANCOVA, a significant main effect of BMI–SDS was observed, $V = 0.53$, $F(2,\ 146) = 82.00$, $p < .001$. Separate univariate ANCOVAs revealed a significant main effect of BMI–SDS for the selection of the current figure, $F(1,\ 147) = 121.72$, $p < .001$, but not for the selection of the ideal figure ($p > .05$). There was a positive relationship between BMI–SDS and current figure choice for both, the ordered ($b = .502$, $SE = .050$, $p < .001$, $r = .636$) and the random scale version ($b = .509$, $SE = .046$, $p < .001$, $r = .676$). There was neither a significant main effect of figure ordering nor a significant interaction effect between figure ordering and BMI–SDS (all $p > .05$). Regarding body dissatisfaction, the results of an ANCOVA revealed a significant main effect of BMI–SDS, $F(1,\ 147) = 124.65$, $p < .001$, with b coefficients indicating a negative relationship between BMI–SDS and body dissatisfaction ($b_{ordered} = -.560$, $SE_{ordered} = .053$, $p_{ordered} < .001$, $r_{ordered} = .654$; $b_{random} = -.523$, $SE_{random} = .052$, $p_{random} < .001$, $r_{random} = .636$). Apart from this, neither a significant main effect of figure ordering nor a significant interaction effect between figure ordering and BMI–SDS were observed (all $p > .05$).

Weight category

Finally, the latter analyses were supplemented using the variable weight category to guarantee a better interpretability of the results concerning weight. Using MANOVA, a significant main effect of weight category was observed, $V = 0.44$, $F(4,\ 292) = 20.56$, $p < .001$. Separate univariate ANOVAs revealed a significant main effect of weight category for the selection of the current figure, $F(2,\ 146) = 50.42$, $p < .001$, but not for the selection of the ideal figure ($p > .05$). Post hoc tests showed that underweight children chose smaller figures to represent their current body than did normal weight children ($M_{under\text{-}normal} = -1.11$, $SE = 0.18$, $p < .001$, $d = 1.580$) and overweight children ($M_{under\text{-}over} = -2.17$, $SE = 0.22$, $p < .001$, $d = 3.115$). Furthermore, normal weight children chose smaller figures to represent their current body than did overweight children ($M_{normal\text{-}over} = -1.06$, $SE = 0.18$, $p < .001$, $d = 1.507$). There was neither a significant main effect of figure ordering nor a significant interaction effect between figure ordering and weight categories (all $p > .05$). Regarding body dissatisfaction, the results of an ANOVA revealed a significant main effect of weight category, $F(2,\ 146) = 37.66$, $p < .001$. Again, post hoc tests showed that underweight children reported less body dissatisfaction than normal weight children ($M_{under\text{-}normal} = 1.11$, $SE = 0.20$, $p < .001$, $d = 1.412$) and overweight children ($M_{under\text{-}over} = 2.11$, $SE = 0.25$, $p < .001$, $d = 2.699$). Furthermore, in normal weight children, ratings of body dissatisfaction were lower than in overweight children ($M_{normal\text{-}over} = 1.00$, $SE = 0.18$, $p < .001$, $d = 1.268$). Finally, the results of an ANOVA revealed neither a significant main effect of figure ordering nor a significant interaction effect between figure ordering and weight category (all $p > .05$).

Criterion and convergent validity

With respect to criterion validity, there was a strong and significant relationship between the selection of current figure and BMI for the scale with ascending figure presentation, $r = .66$, 95% BCa CI [.516, .763], $p < .001$, as well as for the scale with random figure presentation, $r = .68$, 95% BCa CI [.576, .775], $p < .001$. No significant difference between the two correlation coefficients was observed, $t_{diff} = -0.92 < 1.96$, $p > .05$.

Regarding convergent validity, there was a moderate and significant relationship between the EDI–C Body Dissatisfaction subscale and the absolute value of the current–ideal figure discrepancy for the scale with ascending figure presentation, $r = .46$, 95% BCa CI [.335, .575], $p < .001$, and for the scale with random figure presentation, $r = .45$, 95% BCa CI [.323, .582], $p < .001$, with no significant difference between the two correlation coefficients, $t_{diff} = 0.24 < 1.96$, $p > .05$.

Test–retest reliability

For the scale with ascending figure presentation, test–retest reliabilities were very high. Correlations for current figure selection, $\rho = .95$, 95% BCa CI [.885, .989], $p < .001$; ideal figure selection, $\rho = .92$, 95% BCa CI [.823, .992], $p < .001$; and body dissatisfaction, $\rho = .90$, 95% BCa CI [.772, .978], $p < .001$, were significant. Likewise, significant correlations for current figure selection, $\rho = .90$, 95% BCa CI [.775, .987], $p < .001$; ideal figure selection, $\rho = .77$, 95% BCa CI [.544, .935], $p < .001$; and body dissatisfaction, $\rho = .72$, 95% BCa CI [.414, .962], $p < .001$, were observed for the scale with randomized figure presentation. Test–retest reliabilities for this scale version were very high (current figure) and acceptable (ideal figure, body dissatisfaction). All correlations exceeded Nunnally's (1970) criterion of .70 as a minimally acceptable reliability coefficient.

Although there was no significant difference in the test–retest reliabilities of current figure selection, $z_{diff} = 1.80$, $p = .072$, significant differences in test–retest reliabilities were observed for ideal figure selection, $z_{diff} = 2.85$, $p = .004$, and body dissatisfaction, $z_{diff} = 2.82$, $p = .005$. In this regard, correlation coefficients reveal higher accordance when using the scale with ascending figure presentation.

Discussion

Currently, little is known about the psychometric properties of figural rating scales designed to assess body dissatisfaction among children. In particular, the question has been raised whether the ordering of the silhouettes itself has an influence on the results. Previous studies addressing a potential influence of figure arrangement among adolescents and adults have led to inconsistent findings. Therefore, the main goals of our study were to compare the influence of different silhouette orderings (ascending vs. random) on current and ideal figure estimation as well as body dissatisfaction among children 9 to 13 years old. In addition, the reliability and validity of the two presentation modes were assessed.

Comparison of ascending versus random figure ordering

We hypothesized that selections of current and ideal figures as well as the resulting body dissatisfaction are comparable for both presentational methods. Overall, there was a high correspondence between ratings derived from the ascending and randomized presentational methods. No significant differences regarding the selection of current and ideal figures or body dissatisfaction were observed. This observation is in line with results from previous studies addressing adult participants, also finding no influence of the mode of silhouette ordering for figure selection (Doll et al., 2004; Duncan et al., 2005) and body dissatisfaction (Duncan et al., 2005). It might be concluded, therefore, that scales with randomized silhouette ordering would not bring any advantage over scales with ascending figure presentation when assessing body dissatisfaction in children.

In a further step, potential influences from sociodemographic characteristics on figure selections and body dissatisfaction were addressed in an exploratory manner by additional analyses. Especially the role of gender has repeatedly been emphasized in research on body image (see Smolak, 2004). We observed an interaction effect between silhouette ordering and gender such that the current figure rating differed depending on silhouette ordering only among female participants. Whereas girls selected smaller current figures using the ascending scale version compared to the randomized scale version, the boys' estimations remained unaffected by the method of presentation. Although the differences were statistically significant, effect sizes were small. It is not clear, however, whether the ascending or randomized scale version depicts a more realistic self-perception. As Doll et al. (2004) proposed, the ascending order of figures as well as the explicit comparison of adjacent figures might influence a person's perception. In this regard, the figure arrangement within the ordered scale version might emphasize the increase of body sizes and trigger the more pronounced social desirability of thin bodies among girls. In accordance with this assumption, Paxton and Damiano (2017) recently reviewed that the internalization of sociocultural influences like appearance ideals are more pronounced among girls. Contrary to our results, in previous studies applying figural drawing scales with ascending figure presentation, no gender-related differences in the accuracy of self-perception were observed (e.g., Lombardo et al., 2014; Williamson & Delin, 2001). Furthermore, no gender-related differences were observed in previous research on the effect of silhouette ordering (e.g., Duncan et al., 2005; Gardner & Brown, 2010b). It should also be mentioned that in our study no scale-dependent rating was observed for the selection of ideal figures or body dissatisfaction. It is also conceivable that girls might need more orientation regarding the scale direction when judging their body size, which is missing for scales with randomized figure presentation. With respect to age, no age-dependent ratings were observed. However, due to the explorative nature of the complementary analyses and the small effect sizes, no final conclusions should be drawn. Future studies investigating whether gender or age influence the effect of silhouette ordering are warranted.

As observed in previous studies (e.g., $r_{current\ \mathrm{figure\text{-}BMI}} = .69$ reported by Wertheim et al., 2004), the children's ratings of their current body varied with their BMI–SDS respective weight status, supporting the criterion validity of the figure rating scales. Whereas underweight children chose the smallest figures to represent their current body, overweight children selected the largest figures. In line with previous studies (e.g., $r_{\mathrm{body\ dissatisfaction\text{-}BMI}} = |.30|$ reported by Lombardo et al., 2014), ratings of body dissatisfaction also depended on the children's BMI–SDS respective weight status. Underweight children reported the lowest amount of body dissatisfaction and overweight children the highest amount. Effect sizes were high, supporting weight-dependent figure selections and body dissatisfaction. Apart from this observation, the children's weight status did not influence the effect of figure ordering. Taken together, the data strongly suggest that both scale versions are appropriate to measure body satisfaction among girls and boys of different ages and weight.

Psychometric properties

Additionally, the impact of ordered and randomized silhouette presentations on psychometric properties was examined. Both scale versions were shown to be both valid and reliable. High correlations between the figure selection representing the children's current body and the BMI, reflecting good criterion validity, did not vary between the two scale versions ($\rho_{ascending} = .65$, $\rho_{random} = .66$). Furthermore, they were higher than those observed in past studies (e.g., $r = .54$ reported by Coelho et al., 2013; $r = .37$ reported by Collins, 1991; $r = .28–.66$ reported by Lombardo et al., 2014). Hence, a high accuracy of the children's self-perception on the examined scales can be assumed.

Furthermore, acceptable convergent validity was indicated by the moderate correspondence ($r_{ascending} = .46$, $r_{random} = .45$) between judgments of body dissatisfaction using the figural drawing scales and the EDI–C Body Dissatisfaction subscale. The observed correlations were similar to those previously reported among children (e.g., $r = .22–|.64|$ reported by Truby & Paxton, 2002; $r = .40$ reported by Wertheim et al., 2004). Again, no difference between the two scale versions was observed.

Finally, test–retest reliability was tested over a period of 3 days. For the scale with figures presented in ascending order, correlations were higher (ρ between .90 and .95) than those observed in previous studies among children (e.g., $r_{3days} = .59–.71$ reported by Collins, 1991; $r_{14days} = .68–.84$ reported by Wertheim et al., 2004), which might partly be a result of the shorter time period in our study. In contrast, however, correlations for the scale with randomly ordered figures tended to be lower ($\rho = .72–.90$) than those observed in previous studies among adult participants (e.g., $r_{3weeks} = .87$ reported by Gardner et al., 1999). This might be due to the generally lower test–retest reliabilities reported for children compared to adult participants (see Gardner & Brown, 2010a). Among children, both self-perception and idealized body size might be less stable than among adults. Because there are no data on comparable studies with children, comparisons should be drawn carefully.

Interestingly, a difference between the utilized scale versions with regard to test–retest reliabilities was observed, whereby correlation coefficients revealed higher stability of ratings when using the scale with ascending figure presentation. As Gardner et al. (1998) assumed, remembering previously selected figures might require less effort when they are presented in ascending order. However, the observed difference in test–retest reliability coefficients due to figure ordering does not seem to be universal for each rating. The stability of the current body rating was shown to be independent of silhouette ordering. In contrast, a difference was observed for the children's ratings of ideal body and body dissatisfaction, which are probably more vulnerable to external influences and daily fluctuations. Although an anchor effect with a tendency toward the scale's midpoint or to socially desirable answers might lead to a better remembrance of selected figures within the ordered scale version (Gardner et al., 1998), such a reminder does not exist within the scale with randomized silhouette ordering. Hence, children might have had to observe figures within this scale version with higher precision to select an appropriate body size. It can be assumed that this led to more fluctuations, especially for the already less stable ratings of the children's body relative to their desired body size. Future studies should focus on the children's selection behavior and potential differences due to the required rating. Nevertheless, despite the observed differences, test–retest reliabilities for both scale versions were shown to be sufficient to represent psychometrically sound measurements and produce stable ratings.

Strengths and limitations

Several strengths and limitations should be pointed out in this study. First, it can be noted positively that we explicitly compared both presentational methods, extending previous research. Second, we collected objectively measured weight and height data. Third, the applied EDI–C Body Dissatisfaction subscale represents an adequate indicator for the examination of criterion validity.

Besides this, several limitations should be considered in future research on this topic. Although our sample size was sufficient for our main analyses, the investigation of possible moderating influences of age, gender, or weight status was limited. In particular, our study sample included—in line with epidemiological data (Brettschneider, Schienkiewitz, Schmidt, Ellert, & Kurth, 2017)—only a small proportion of underweight and overweight children. As mentioned before, future research will be required to address potential influences of age, gender, and weight status on the effect of silhouette ordering as well as on data regarding validity and reliability. In addition, we decided not to ask the parents to fill in questionnaires. Therefore, we have no data on the socioeconomic background of our sample. Furthermore, by investigating the influence of the silhouette ordering, we did not examine all the different ordering possibilities of the scales. Another point concerns the silhouette scale we applied: Although the scale developed by Collins (1991) is probably the most commonly used (Smolak, 2004), several new scales have since been developed, applying photographic figures (e.g., Truby & Paxton, 2002). Meers, Koball, Oehlhof, Laurene, and Musher-Eizenman (2011) argued that those further developed scales might represent a better measurement tool in the assessment of weight-related bias in preschoolers. Although the degree of figure realism is not assumed to influence results on our main research question concerning the influence of figure ordering, future studies might consider scales with photographic figures. Finally, the 10-min break between the two scale versions might have been too short to ensure a sufficient distraction. Future studies should take these considerations into account.

Conclusions

The data strongly suggest that ascending figure presentation does not cause a response bias in measuring body satisfaction among children, and that these scales are suitable for children of different ages and weight status. Initial tentative indications show that girls might be more affected by the silhouette ordering than boys, so that gender differences would warrant further investigation.

Furthermore, the ascending and randomized scale versions were shown to be valid as well as reliable. Although convergent and criterion validity did not vary between the two scale versions, correlation coefficients revealed a higher stability of body dissatisfaction ratings when using the scale with ascending figure presentation. The scale with randomized silhouette ordering might offer more space for external influences and fluctuations.

Due to the easy and fast application, silhouette scales represent an economical assessment tool to determine body dissatisfaction in preadolescent participants.

Acknowledgments

Special thanks to the participants. Thanks to the teachers of the participating elementary schools for supporting the study. Thanks to Susannah Motter and Christina Schrezenmeir for supporting the collection of data. Jana Zitzmann conceived, designed, and coordinated the study, performed the statistical analysis, participated in the interpretation of the data, and wrote and revised the manuscript. Petra Warschburger conceived and designed the study, participated in the interpretation of the data, and participated in writing and revising the manuscript. All authors read and approved the final article.

References

Brettschneider, A.-K., Schienkiewitz, A., Schmidt, S., Ellert, U., & Kurth, B.-M. (2017). Updated prevalence rates of overweight and obesity in 4- to 10-year-old children in Germany: Results from the telephone-based KiGGS wave 1 after correction for bias in parental reports. *European Journal of Pediatrics*, *176*, 547–551. doi:10.1007/s00431-017-2861-8

Chen, P. Y., & Popovich, P. M. (2002). *Correlation: Parametric and nonparametric measures. Sage University Papers series on quantitative applications in the social sciences* (pp. 7–139). Thousand Oaks, CA: Sage.

Coelho, E. M., Padez, C., Moreira, P., Rosado, V., & Mourão-Carvalhal, I. (2013). BMI and self-perceived body shape in Portuguese children. *Revista de Psicología del Deporte*, *22*, 371–376.

Collins, M. E. (1991). Body figure perceptions and preferences among preadolescent children. *International Journal of Eating Disorders*, *10*, 199–208. doi:10.1002/1098-108X(199103)10:2 < 199::AID-EAT2260100209 > 3.0.CO;2-D

Dion, J., Hains, J., Vachon, P., Plouffe, J., Laberge, L., Perron, M., … Leone, M. (2016). Correlates of body dissatisfaction in children. *Journal of Pediatrics*, *171*, 202–207. doi:10.1016/j.jpeds.2015.12.045

Doll, M., Ball, G. D. C., & Willows, N. D. (2004). Rating of figures used for body image assessment varies depending on the method of figure presentation. *International Journal of Eating Disorders*, *35*, 109–114. doi:10.1002/eat.10233

Duncan, M. J., Dodd, L. J., & Al-Nakeeb, Y. (2005). The impact of silhouette randomization on the results of figure rating scales. *Measurement in Physical Education and Exercise Science*, *9*, 61–66. doi:10.1207/s15327841mpee0901_5

Fallon, A. E., & Rozin, P. (1985). Sex differences in perceptions of desirable body shape. *Journal of Abnormal Psychology*, *94*, 102–105.

Field, A. (2015). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock'n'roll* (4th ed.). Los Angeles: Sage.

Furnham, A., & Alibhai, N. (1983). Cross-cultural differences in the perception of female body shapes. *Psychological Medicine*, *13*, 829–837.

Gardner, R. M. (1996). Methodological issues in assessment of the perceptual component of body image disturbance. *British Journal of Psychology*, *87*, 327–337. doi:10.1111/j.2044-8295.1996.tb02593.x

Gardner, R. M., & Brown, D. L. (2010a). Body image assessment: A review of figural drawing scales. *Personality and Individual Differences*, *48*, 107–111. doi:10.1016/j.paid.2009.08.017

Gardner, R. M., & Brown, D. L. (2010b). Comparison of video distortion and figural drawing scale for measuring and predicting body image dissatisfaction and distortion. *Personality and Individual Differences*, *49*, 794–798.

Gardner, R. M., & Brown, D. L. (2011). Method of presentation and sex differences when using a revised figural drawing scale to measure body size estimation and dissatisfaction. *Perceptual and Motor Skills*, *113*, 739–750. doi:10.2466/07.17.27.PMS.113.6.739-750

Gardner, R. M., Espinoza, T., Urrutia, R., Morrell, J., & Gallegos, V. (1989). Effect of hunger on body image in the obese: Feeling full, feeling fat. *Psychiatric Forum*, 15, 57–62.

Gardner, R. M., Friedman, B. N., & Jackson, N. A. (1998). Methodological concerns when using silhouettes to measure body image. *Perceptual and Motor Skills*, *86*, 387–395. doi:10.2466/pms.1998.86.2.387

Gardner, R. M., Jappe, L. M., & Gardner, L. (2009). Development and validation of a new figural drawing scale for body-image assessment: The BIAS-BD. *Journal of Clinical Psychology*, *65*, 113–122. doi:10.1002/jclp.20526

Gardner, R. M., Martínez, R., & Espinoza, T. (1987). Psychophysical measurement of body image of self and others in obese subjects. *Journal of Social Behavior and Personality*, *2*, 205–218.

Gardner, R. M., Martínez, R., Espinoza, T., & Gallegos, V. (1988). Distortion of body image in the obese: A sensory phenomenon. *Psychological Medicine*, *18*, 633–641.

Gardner, R. M., Stark, K., Jackson, N. A., & Friedman, B. N. (1999). Development and validation of two new scales for assessment of body-image. *Perceptual and Motor Skills*, *89*, 981–993. doi:10.2466/pms.1999.89.3.981

Garner, D. M. (1991). *The Eating Disorder Inventory-C*. Lutz, FL: Psychological Assessment Resources, Inc.

Garner, D. M., & Garfinkel, P. E. (1981). Body image in anorexia nervosa: Measurement, theory and clinical implications. *International Journal of Psychiatry in Medicine*, *11*, 263–284.

Heck-Möhling, R., Reinhard, J., & Boehle, J. (1993). *Konzentrationstest für 3. und 4. Klassen (KT 3/4)*. Weinheim: Beltz.

Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., Geller, F., Geiß, H. C., Hesse, V., … Hebebrand, J. (2001). Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschrift Kinderheilkunde*, *149*, 807–818. doi:10.1007/s001120170107

Lombardo, C., Battagliese, G., Pezzuti, L., & Lucidi, F. (2014). Validity of a figure rating scale assessing body size perception in school-age children. *Eating and Weight Disorders*, *19*, 329–336. doi:10.1007/s40519-013-0085-0

Meers, M. R., Koball, A. M., Oehlhof, M. W., Laurene, K. R., & Musher-Eizenman, D. R. (2011). Assessing anti-fat bias in preschoolers: A comparison of a computer generated line-drawn figure array and photographic figure array. *Body Image*, *8*, 293–296.

Nicholls, M. E. R., Orr, C. A., Okubo, M., & Loftus, A. (2006). Satisfaction guaranteed: The effect of spatial biases on responses to likert scales. *Psychological Science*, *17*, 1027–1028. doi:10.1111/j.1467-9280.2006.01822.x

Nunnally, J. C. (1970). *Introduction to Psychological Measurement*. New York: McGraw-Hill.

Paludo, A. C., Pelegrini, A., Grespan, F., Caldeira, A. S., Madureira, A. S., & Serassuelo Junior, H. (2011). Insatisfação com a imagem corporal em adolescentes: Prevalência e associação com o estado nutricional. *ConScientiae Saúde*, *10*, 143–149.

Paxton, S. J., & Damiano, S. R. (2017). The development of body image and weight bias in childhood. *Advances in Child Development and Behavior*, *52*, 269–298. doi:10.1016/bs.acdb.2016.10.006

Rand, C. S. W., & Wright, B. A. (2001). Thinner females and heavier males: Who says? A comparison of female to male ideal body sizes across a wide age span. *International Journal of Eating Disorders*, *29*, 45–50. doi:10.1002/1098-108X(200101)29:1 < 45::AID-EAT7 > 3.0.CO;2-I

Sala, L., Mirabel-Sarron, C., Pham-Scottez, A., Blanchet, A., Rouillon, F., & Gorwood, P. (2012). Body dissatisfaction is improved but the ideal silhouette is unchanged during weight recovery in anorexia nervosa female inpatients. *Eating and Weight Disorders*, *17*, 109–115.

Sherman, D. K., Iacono, W. G., & Donnelly, J. M. (1995). Development and validation of body rating scales for adolescent females. *International Journal of Eating Disorders*, *18*, 327–333.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428. doi:10.1037/0033-2909.86.2.420

Smolak, L. (2004). Body image in children and adolescents: Where do we go from here? *Body Image*, *1*, 15–28. doi:10.1016/S1740-1445(03)00008-1

Song, A. Y., Rubin, J. P., Thomas, V., Dudas, J. R., Marra, K. G., & Fernstrom, M. H. (2006). Body image and quality of life in post massive weight loss body contouring patients. *Obesity*, *14*, 1626–1636.

Stunkard, A. J., Sorensen, T. I., & Schulsinger, F. (1983). Use of the Danish adoption register for the study of obesity and thinness. In S. Kety, L. P. Rowland, R. L. Sidman, & S. W. Matthysse (Eds.), *The genetics of neurological and psychiatric disorders* (pp. 115–120). New York: Raven Press.

Tatangelo, G., McCabe, M., Mellor, D., & Mealey, A. (2016). A systematic review of body dissatisfaction and sociocultural messages related to the body among preschool children. *Body Image*, *18*, 86–95. doi:10.1016/j.bodyim.2016.06.003

Thiels, C., Salbach-Andrae, H., Bender, C., & Garner, D. M. (2011). EDI-C – Eating Disorder Inventory – C (deutsche Version). In C. Barkmann, M. Schulte-Markwort, & E. Brähler (Eds.), *Ratingskalen zur Diagnostik klinisch-psychiatrischer Syndrome des Kindes- und Jugendalters* (pp. 218–222). Göttingen: Hogrefe.

Thompson, J. K. (1995). Assessment of body image. In D. B. Allison (Ed.), *Handbook of assessment methods for eating behaviors and weight-related problems: Measures, theory and research*. Thousand Oaks, CA: Sage.

Thompson, J. K., & Altabe, M. N. (1991). Psychometric qualities of the figure rating scale. *International Journal of Eating Disorders*, *10*, 615–619. doi:10.1002/1098-108X(199109)10:5 < 615::AID-EAT2260100514 > 3.0.CO;2-K

Thompson, M. A., & Gray, J. J. (1995). Development and validation of a new body-image assessment scale. *Journal of Personality Assessment*, *64*, 258–269. doi:10.1207/s15327752jpa6402_6

Thompson, J. K., Heinberg, L. J., Altabe, M. N., & Tantleff-Dunn, S. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance*. Washington: American Psychological Association.

Thurfjell, B., Edlund, B., Arinell, H., Hagglof, B., & Engstrom, I. (2003). Psychometric properties of eating disorder inventory for children (EDI-C) in Swedish girls with and without a known eating disorder. *Eating and Weight Disorders*, *8*, 296–303.

Tremblay, L., Lovsin, T., Zecevic, C., & Lariviere, M. (2011). Perceptions of self in 3-5-year-old children: A preliminary investigation into the early emergence of body dissatisfaction. *Body Image*, *8*, 287–292. doi:10.1016/j.bodyim.2011.04.004

Truby, H., & Paxton, S. J. (2002). Development of the children's body image scale. *British Journal of Clinical Psychology*, *41*, 185–203.

Walker, L. L. M., Gately, P. J., Bewick, B. M., & Hill, A. J. (2003). Children's weight-loss camps: Psychological benefit or jeopardy? *International Journal of Obesity*, *27*, 748–754.

Wertheim, E. H., Paxton, S. J., & Tilgner, L. (2004). Test-retest reliability and construct validity of Contour Drawing Rating Scale scores in a sample of early adolescent girls. *Body Image*, *1*, 199–205. doi:10.1016/S1740-1445(03)00024-X

Williamson, S., & Delin, C. (2001). Young children's figural selections: Accuracy of reporting and body size dissatisfaction. *International Journal of Eating Disorders*, *29*, 80–84.