Obesity Facts

Research Article

Obes Facts 2024;17:329–337
DOI: 10.1159/000537689

Received: November 9, 2023
Accepted: February 1, 2024
Published online: February 13, 2024

Psychometric Properties of the WBIS/M in a Representative Prebariatric Sample: Evidence for an Improved 10-Item Version

Simon Schraven[a] Claudia Hübner[a] Janina Eichler[a] Thomas Mansfeld[b] Johannes Sander[c] Florian Seyfried[d] Stefan Kaiser[e] Arne Dietrich[f] Ricarda Schmidt[a] Anja Hilbert[a]

[a]Integrated Research and Treatment Center Adiposity Diseases, Behavioral Medicine Research Unit, Department of Psychosomatic Medicine and Psychotherapy, University of Leipzig Medical Center, Leipzig, Germany; [b]Department of General Surgery, Asklepios Clinic Hamburg, Hamburg, Germany; [c]Schön Klinik Hamburg Eilbek, Obesity Clinic Hamburg, Hamburg, Germany; [d]Department of General, Visceral, Transplant, Vascular and Pediatric Surgery, University Hospital Würzburg, Würzburg, Germany; [e]Department of Visceral, Pediatric and Vascular Surgery, Hospital Konstanz, Konstanz, Germany; [f]Department of Surgery, Clinic for Visceral, Transplantation, Thoracic and Vascular Surgery, University Hospital Leipzig, Leipzig, Germany

Keywords
Confirmatory factor analysis · Germany · Prebariatric samples · Psychometrics · WBIS

Abstract
Introduction: The Weight Bias Internalization Scale and the Modified Weight Bias Internalization Scale are well-established self-report questionnaires for assessing weight bias internalization, which is widespread among bariatric patients. However, among this group, psychometric properties of the Weight Bias Internalization Scale have only been examined in small samples showing unsatisfactory model fit and have not been explored for the modified questionnaire. ***Methods:*** This study psychometrically evaluated and compared the Weight Bias Internalization Scale and Modified Weight Bias Internalization Scale in a large sample of prebariatric patients (N = 825, mean age = 46.75 years, SD = 11.55) regarding item characteristics, model fit to unidimensionality, reliability, construct validity, and measurement invariance. ***Results:*** Item 4 of both questionnaires showed low corrected item-total correlations (<0.40) and was therefore removed from the scales. The new 10-item versions showed improved item characteristics, internal consistency, model fit to unidimensionality, and convergent and divergent validity when compared to the 11-item versions. The best psychometric properties were found for the 10-item version of the Modified Weight Bias Internalization Scale. ***Conclusion:*** The 10-item version of the Modified Weight Bias Internalization Scale surpasses the other versions studied in all psychometric properties. Therefore, it should be used in prebariatric patients to detect weight bias internalization and provide them with psychological interventions that could improve bariatric surgery outcomes.

Introduction

Weight bias describes negative prejudices and stereotypes regarding a person's body weight, including attributions such as laziness, lack of willpower, or

karger@karger.com
www.karger.com/ofa

Correspondence to:
Anja Hilbert, anja.hilbert@medizin.uni-leipzig.de

moral character. Relatedly, individuals with overweight (body mass index [BMI]; 25.0 ≤ BMI ≤29.9 kg/m^2) and obesity (BMI ≥30.0 kg/m^2) experience weight-related stigmatization and discrimination in many areas of life (e.g., employment, health care, and education) [1]. Additionally, individuals with overweight and obesity tend to apply public weight bias to themselves, termed *weight bias internalization* (WBI), which is associated with self-loathing and low self-esteem [2]. Furthermore, a meta-analysis showed WBI to be positively correlated with depression and anxiety [3]. Finally, a longitudinal study [4] as well as a meta-analysis [5] showed WBI and these psychopathologies to be negatively correlated with weight loss after bariatric surgery, which is the most efficacious long-term weight loss method in patients with severe obesity (BMI ≥40.0 kg/m^2 or BMI ≥35.0 kg/m^2 with obesity-related comorbidity) [6, 7]. Reliable and valid means of assessing WBI in this patient population are therefore important.

Durso and Latner [2] developed the Weight Bias Internalization Scale (WBIS), an 11-item self-report questionnaire measuring to which degree individuals apply negative obesity-related attributions to themselves (e.g., "I don't feel that I deserve to have a really fulfilling social life, as long as I'm overweight") [2]. In addition, a Modified WBIS (WBIS-M) that assesses WBI using weight-neutral language (e.g., "I don't feel that I deserve to have a really fulfilling social life, as long as I'm my weight") [8] among various weight categories has been derived from the WBIS. A systematic review collating psychometric properties of WBIS/-M [9] in population-based samples of adults (148 ≤ N ≤ 279) and samples of adults with overweight and obesity (90 ≤ N ≤ 1,092) underlined "sufficient" evidence for internal consistency and confirmed construct validity of measures but revealed inconsistencies regarding the assumed unidimensionality [2]. In bariatric surgery, only three psychometric studies with relatively small sample sizes of prebariatric adolescents (N = 57, 14 ≤ age ≤18 years) [10] and adults (N = 78/253) [11, 12] have been conducted so far. In these studies, the WBIS showed good to excellent internal consistency (Cronbach's α = 0.84/0.92) [10, 11] and large-sized corrected item-total correlations (r_{it} = 0.50/0.90) except of item 1 ("feeling competent," r_{it} = 0.34/0.37) [10, 12]. Consequently, item 1 was removed resulting in a modified 10-item version. Additionally, psychometric studies supported the convergent validity of the WBIS, indicated by large-sized correlations with another self-report questionnaire on WBI [11] and clinically related constructs such as depression (r = 0.52/0.58) [10, 12] and anxiety (r = 0.57) [12]. However, the unidimensionality of the WBIS was hardly supported, but slightly improved when item 1 was removed [12]. Altogether, the investigated prebariatric samples were rather small-sized and several psychometric properties about the WBIS were not examined (e.g., McDonald's Ω, divergent validity). Strikingly, the WBIS-M has not been evaluated among prebariatric samples so far. Hence, this study is the first assessing and comparing psychometric properties (item characteristics, model fit to unidimensionality, reliability, convergent and divergent validity) of the WBIS and WBIS-M in a large prebariatric sample.

Methods

Study Design and Participants

The present study is embedded in the multicenter Psychosocial Registry for Obesity Surgery (PRAC). The ongoing survey prospectively examines psychosocial aspects of patients undergoing bariatric surgery who were recruited at six German bariatric surgery centers. The data were collected independently from clinical treatment and were not shared with the surgical team. Inclusion criteria for PRAC were planned bariatric surgery, whereas exclusion criteria were noncompliance and insufficient German language ability. The present study included data of participants with BMI ≥35.0 kg/m^2 and age ≥18 years, who planned bariatric surgery and completed the WBIS/-M between March 1, 2012, and August 31, 2022. The WBIS-M, assessing WBI among various weight categories by using weight-neutral language [7], replaced the WBIS in April 2015 in the PRAC study. Thus, all statistical analyses were performed separately for participants completing the WBIS (n = 325) and the WBIS-M (n = 500). The subsample characteristics are presented in Table 1.

Measures

WBIS and WBIS-M

The WBIS [2] contains 11 items that are rated on a 7-point Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*. The present study used the German versions of the WBIS/-M, which were translated into German and controlled by a back-translation by a licensed translator [13]. A mean score, with higher scores indicating higher WBI, was computed.

Measures for Convergent Validation and Divergent Validation

Convergent validity was determined by the weight-related items of the Perception of Teasing Scale (POTS), assessing the frequency of perceived weight-related childhood teasing [Hilbert, unpublished results; 14]; the Eating Disorder Examination-Questionnaire (EDE-Q) subscales eating, weight, and shape concern [15, 16], assessing eating disorder psychopathology; the Patient Health Questionnaire (PHQ-9) [17],

Table 1. Sample characteristics

Baseline characteristics	WBIS[a] (n = 325)		WBIS-M[b] (n = 500)		T test for differences		
	M or n	SD or %	M or n	SD or %	df	t	p value
Sex[c], women	224	68.92	330	66.00	823	−8.73	0.383
Age, years	44.75	10.89	47.93	11.71	823	3.96	<0.001
BMI, kg/m^2	49.27	7.65	48.32	8.01	823	−1.69	0.091
Weight status							
Obesity class 2[d]	24	7.38	78	15.60			
Obesity class 3[e]	301	92.62	422	84.40			

M, mean; SD, standard deviation; % from valid cases n; df, degree of freedom; t, t value; p, p value; BMI, body mass index. [a]WBIS, Weight Bias Internalization Scale; [b]WBIS-M, Modified Weight Bias Internalization Scale; the WBIS was used until March 2015, and the WBIS-M was used from April 2015; [c]self-reported sex; [d]obesity class 2 (35.0 ≤ BMI ≤ 39.9 kg/m^2); [e]obesity class 3 (BMI ≥40.0 kg/m^2).

assessing depression; the Impact of Weight on Quality of Life-Lite (IWQOL-Lite) [18, 19], assessing body weight-related quality of life; and the General Self-Efficacy Scale (GSES) [20], assessing self-efficacy. Divergent validity was determined by the EDE-Q subscale restraint [15], assessing attempts to dietary restriction. Among both subsamples, the internal consistency of the POTS (α/ω_{WBIS} = 0.96/0.96, $\alpha/\omega_{WBIS\text{-}M}$ = 0.97/0.97), the EDE-Q (α/ω_{WBIS} = 0.85/0.84, $\alpha/\omega_{WBIS\text{-}M}$ = 0.88/0.87), the PHQ-9 (α/ω_{WBIS} = 0.85/0.86, $\alpha/\omega_{WBIS\text{-}M}$ = 0.85/0.86), the IWQOL-Lite (α/ω_{WBIS} = 0.94/0.94, $\alpha/\omega_{WBIS\text{-}M}$ = 0.96/0.95), and the GSES (α/ω_{WBIS} = 0.93/0.93, $\alpha/\omega_{WBIS\text{-}M}$ = 0.93/0.93) was good to excellent. The measures for validation are described in detail in the online supplement S1 (for all online suppl. material, see https://doi.org/10.1159/000537689).

Statistical Methodology

SPSS for Windows, version 27.0 [21]; AMOS, version 29.0 [22]; and ARTool 2.0 [23] were used for data analysis, and a two-tailed α of 0.05 was applied.

Item Analyses

Missing data were analyzed as the percentage of missing item responses per item and replaced by mean imputation for items with <5% of missing item responses [24]. Corrected item-total correlations (r_{it}) were calculated and interpreted as sufficient if the r_{it} was ≥0.30 [25]. The 11-item version and versions after removing items with r_{it}<0.30 were compared regarding all following psychometric properties. The absolute values of the item skewness and kurtosis were calculated, and the items were interpreted as normally (absolute value = 0.00), slightly non-normally (absolute value <1.00), moderately non-normally (1.00 ≤ absolute value ≤2.30), and severely non-normally distributed (absolute value >2.30) [26], respectively. Additionally, item difficulty was calculated as p_m = sum of item scores/(N × maximal item score) and interpreted as more difficult if closer to 0.00 and easier if closer to 1.00 [27]. WBIS/-M mean scores were computed and tested for normality by Shapiro-Wilk test. Given non-normality outliers (standard deviation [*SD*] of WBIS/-M mean score beyond ±3) were excluded listwise [28] and WBIS/-M mean scores were tested for differences by Mann-Whitney U test.

Factor Analyses

A confirmatory factor analysis using AMOS [22] examined the hypothesized unidimensionality of the WBIS/-M [2, 29]. Multivariate normality was examined by the Mardia test and rejected if the absolute value of the critical ratio was >1.96 [30]. If non-normality was given, the Bollen-Stine bootstrap method was applied to investigate the adequacy of fit [31]. To determine model fit, the minimum discrepancy divided by its degrees of freedom (CMIN/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR) were calculated. An acceptable model fit was indicated by 2.00 < CMIN/df ≤3.00, 0.95 ≤ CFI <0.97, 0.90 ≤ TLI <0.95, 0.06 ≤ RMSEA ≤0.10, whereas a good model fit was indicated by CMIN/df ≤2.00, CFI ≥0.97, TLI ≥0.95, RMSEA <0.06 [27], and SRMR ≤0.08 [32]. Factor loadings <0.40 were considered as random [33].

Reliability

Cronbach's α and McDonald's ω were computed for internal consistency and interpreted as acceptable if 0.70 ≤ α/ω <0.80, good if 0.80 ≤ α/ω <0.90, and excellent if α/ω ≥0.90 [34]. Homogeneity was examined by mean inter-item correlations and interpreted as small if 0.10 ≤ r <0.30, medium if 0.30 ≤ r <0.50, and large if r ≥0.50 [35].

Convergent and Divergent Validity

Spearman rank correlation coefficients of the WBIS/-M mean scores with measures for convergent validation (POTS; EDE-Q eating concern, weight concern, and shape concern; PHQ-9; IWQOL-Lite; GSES) and divergent validation (EDE-Q restraint) were calculated, and the absolute values were interpreted according to Cohen [35].

Distributions of WBIS/-M Mean Scores

An analysis of variance (ANOVA) of WBIS/-M mean scores regarding self-determined sex (female or male), age (≤24, 25–34, 35–44, 45–54, 55–64, ≥64 years), and weight status (obesity class 2, obesity class 3) was performed. To this end, WBIS/-M mean scores were tested for homoscedasticity by the Brown-Forsythe test [36]. If non-normality and homoscedasticity were given, a nonparametric aligned rank transform (ART) [37] ANOVA was performed using the ARTool 2.0 [23].

Results

Item Analyses

Item characteristics and WBIS/-M mean scores are displayed in Table 2. Missing item responses occurred rarely (0.03–0.04%) regarding both self-report questionnaires. Corrected item-total correlations were medium- to large-sized ($0.39 \leq r_{it} \leq 0.81$) among both questionnaires, with one exception: item 4's item-total correlations ("weight-change desire," $r_{it} = 0.20/0.23$) fell below the threshold of 0.30, leading to analyses of the original 11-item and 10-item versions of the WBIS/-M after removing item 4. The r_{it} was similar comparing the 10-item and the 11-item versions of the WBIS/-M but was mostly higher for the WBIS-M than for the WBIS. Only the r_{it} of item 1 was equal (11-item version) or higher (10-item version) for the WBIS than for the WBIS-M. Skewness and kurtosis indicated slight to moderate non-normality for most items (skewness and kurtosis <2.30). However, skewness and kurtosis of item 9 of the WBIS (skewness = 2.04; kurtosis = 4.38) and item 4 of both questionnaires (WBIS, skewness = −3.54, kurtosis = 17.46; WBIS-M, skewness = −3.31; kurtosis = 14.75) indicated severe non-normality. The item difficulty indices ranged from extremely difficult ($p_m = 0.26/0.30$, item 9) to very easy ($p_m = 0.93/0.96$, item 4) and were indicative for greater difficulty of the WBIS, except for item 1 and 9, which were more difficult regarding the WBIS-M. The Shapiro-Wilk test of the WBIS/-M mean scores indicated non-normality for both questionnaires ($p < 0.05$), and the WBIS/-M mean scores differed significantly from each other regarding both the 11-item and 10-item versions (11-item versions, $U = 71{,}272.50$, $Z = -2.98$, $p < 0.05$; 10-item versions, $U = 71{,}720.00$, $Z = -2.85$, $p < 0.05$).

Factor Analyses

Mardia test indicated multivariate non-normality for the 11-item and 10-item versions of both subsamples (13.20 ≤ critical ratio ≤23.29). Hence, the Bollen-Stine bootstrap method with $n = 2{,}000$ bootstraps was used prior to the confirmatory factor analysis. The model fit of the WBIS was good regarding SRMR (11-item version, 0.07; 10-item version, 0.06) and acceptable regarding RMSEA (both versions, 0.10, 90% confidence interval [CI]: 0.09–0.12) but was unacceptable regarding CMIN/df (11-item version, 4.44; 10-item version, 4.15), TLI (11-item version, 0.84; 10-item version, 0.88), and CFI (11-item version, 0.86; 10-item version, 0.91).

The model fit of the WBIS-M was good regarding SRMR (both versions, 0.05), acceptable regarding TLI (11-item version, 0.91; 10-item version, 0.92) and RMSEA (11-item version, 0.09, 90% CI: 0.08–0.10; 10-item version, 0.08, 90% CI: 0.07–0.10), but unacceptable regarding CMIN/*df* (11-item version, 4.85; 10-item version, 5.11) and CFI (11-item version, 0.92; 10-item version, 0.94). Importantly, item 4 of the 11-item versions of the WBIS/-M showed factor loadings <0.40, while the 10-item versions only showed factor loadings >0.40 (WBIS, 0.42–0.85; WBIS-M, 0.42–0.88). Generally, the WBIS-M displayed higher factor loadings than the WBIS, except for items 1 and 9. Detailed factor loadings are presented in the online supplement (online suppl. Table S1).

Reliability

The internal consistency was good for the WBIS (11-item version, Cronbach's α = 0.85, McDonalds ω = 0.85; 10-item version, Cronbach's α = 0.86, McDonalds ω = 0.86) and the WBIS-M (11-item version, Cronbach's α = 0.87, McDonalds ω = 0.88; 10-item version, Cronbach's α = 0.88, McDonalds ω = 0.89). The mean inter-item correlations were smaller for the WBIS (11-item version, 0.33, SD = 0.14; 10-item version, 0.38, SD = 0.14) than for the WBIS-M (11-item version, 0.38, SD = 0.17; 10-item version, 0.43, SD = 0.14) and smaller for the 11-item than for the 10-item version regarding both questionnaires.

Convergent and Divergent Validity

The Spearman rank correlation coefficients with the measures for convergent and divergent validation are displayed in Table 3. The WBIS and the WBIS-M showed convergent validity as indicated by medium- to large-sized correlations with the EDE-Q subscales eating, shape, and weight concern, the PHQ-9, the IWQOL-Lite, and the GSES. However, the WBIS showed only a small-sized correlation with the POTS, while a medium-sized correlation was found for the WBIS-M. Divergent validity was shown for both self-report questionnaires by small-

Table 2. Item characteristics of the WBIS[a] (n = 325) and the WBIS-M[a] (n = 500)

Item or variable	M	SD	$r_{it\text{-}11}$	$r_{it\text{-}10}$[b]	Skewness	Kurtosis	p_m	Missing item responses, n (%)
1. Feeling competent								
WBIS	4.95	1.95	0.44	0.45	−0.67	−0.85	0.71	0 (0.00)
WBIS-M	5.23	1.76	0.42	0.43	−0.96	−0.07	0.75	0 (0.00)
2. Feeling attractive								
WBIS	5.45	1.59	0.44	0.43	−1.16	0.70	0.78	2 (0.63)
WBIS-M	5.06	1.70	0.52	0.51	−0.89	−0.18	0.72	0 (0.00)
3. Anxiousness about weight								
WBIS	4.93	1.86	0.57	0.56	−0.77	−0.51	0.70	0 (0.00)
WBIS-M	4.54	1.92	0.70	0.70	−0.51	−0.96	0.65	0 (0.00)
4. Weight-change desire								
WBIS	6.69	0.74	0.20		−3.54	17.46	0.96	1 (0.32)
WBIS-M	6.54	0.90	0.24		−3.31	14.75	0.93	2 (0.40)
5. Feeling depressed								
WBIS	5.42	1.71	0.70	0.69	−1.16	0.52	0.77	0 (0.00)
WBIS-M	5.06	1.87	0.73	0.73	−0.87	−0.38	0.72	2 (0.40)
6. Self-hate								
WBIS	4.63	2.06	0.68	0.68	−0.42	−1.10	0.66	0 (0.00)
WBIS-M	4.16	2.24	0.77	0.77	−0.20	−1.46	0.59	0 (0.00)
7. Self-value								
WBIS	4.91	1.93	0.78	0.78	−0.72	−0.64	0.70	0 (0.00)
WBIS-M	4.38	2.10	0.81	0.81	−0.38	−1.24	0.63	2 (0.40)
8. Deserving no socially fulfilling life								
WBIS	2.95	2.08	0.53	0.54	0.69	−0.91	0.42	1 (0.32)
WBIS-M	2.78	1.97	0.55	0.56	0.73	−0.80	0.40	2 (0.40)
9. Feeling okay (r)								
WBIS	1.80	1.29	0.39	0.39	2.04	4.38	0.26	1 (0.32)
WBIS-M	2.07	1.43	0.42	0.41	1.48	1.63	0.30	1 (0.20)
10. Not being true self								
WBIS	4.02	2.12	0.58	0.59	−0.02	−1.32	0.57	2 (0.63)
WBIS-M	3.76	2.04	0.63	0.63	0.11	−1.27	0.54	1 (0.20)
11. Not meriting to be dated								
WBIS	4.76	1.96	0.57	0.57	−0.50	−0.97	0.68	0 (0.00)
WBIS-M	4.34	2.00	0.65	0.65	−0.30	−1.14	0.62	4 (0.80)
Total scale 11-item version								
WBIS	4.59	0.90			−0.35	−0.21		7 (0.03)
WBIS-M	4.36	1.03			−0.35	−0.39		14 (0.04)
Total scale 10-item version								
WBIS	4.38	0.97			−0.31	−0.38		6 (0.03)
WBIS-M	4.27	1.18			−0.30	−0.65		12 (0.03)

M, mean; SD, standard deviation; p_m, item difficulty; $r_{it\text{-}11}$, corrected item-total correlation of the 11-item version of the WBIS/-M; $r_{it\text{-}10}$, corrected item-total correlation of the 10-item version of the WBIS/-M; (r), reverse scored. [a]The WBIS was used until March 2015, and the WBIS-M was used from April 2015; [b]in the 10-item version, item 4 ("weight-change desire") was removed.

Table 3. Spearman rank correlation coefficients (r) of the *WBIS*[a] (n = 325) and the *WBIS-M*[a] (n = 500) with measures for convergent and divergent validation

Measures for validation	*r* of the WBIS		*r* of the WBIS-M	
	11-item version	10-item version[b]	11-item version	10-item version[b]
Convergent validity				
POTS[c]	0.25**	0.26**	0.39**	0.39**
EDE-Q eating concern[d]	0.51**	0.51**	0.59**	0.59**
EDE-Q shape concern[d]	0.62**	0.62**	0.66**	0.65**
EDE-Q weight concern[d]	0.59**	0.59**	0.64**	0.63**
IWQOL-Lite[e]	−0.61**	−0.62**	−0.63**	−0.63**
GSES[f]	−0.30**	−0.31**	−0.40**	−0.40**
PHQ-9[g]	0.52**	0.52**	0.54**	0.54**
Divergent validity				
EDE-Q restraint[d]	0.13*[h]	0.13*[h]	0.14**	0.13**

r, Spearman rank correlation coefficients. *$p < 0.05$. **$p < 0.001$. [a]The WBIS was used until March 2015, and the WBIS-M was used from April 2015; [b]in the 10-item version, item 4 ("weight-change desire") was removed; [c]POTS, Perception of Teasing Scale mean score (1–5[#], [#]showing less favorable scores); [d]EDE-Q, Eating Disorder Examination-Questionnaire subscale mean score (0–6[#]); [e]IWQOL-Lite, Impact of Weight on Quality of Life-Lite transformed sum score (0[#]–100); [f]GSES, General Self-Efficacy Scale sum score (4[#]–40); [g]PHQ-9, Patient Health Questionnaire 9 sum score (0–27[#]); [h]$p = 0.02$.

sized correlations with the EDE-Q subscale restraint. Overall, the WBIS revealed smaller validity indicators than the WBIS-M, with few differences between the 11-item and the 10-item versions.

Distributions of WBIS/-M Mean Scores

The residuals of WBIS/-M mean scores showed non-normality as examined with the Shapiro-Wilk test ($p < 0.001$), but homoscedasticity was given as examined with the Brown-Forsythe test ($p > 0.05$). Consequently, an ART-ANOVA was performed, which did not indicate any significant effects of sex, age, weight status, or their interactions on the WBIS/-M mean scores (all ps > 0.05). The detailed results of the ART-ANOVA are presented in the online supplement (online suppl. Table S2).

Discussion

This is the first psychometric investigation of the WBIS/-M in a large prebariatric sample. Because item 4 ("weight-change desire") of the WBIS/-M showed low corrected item-total correlations, the 10-item versions (without item 4) of the WBIS/-M have been evaluated, additionally to the original 11-item versions. The 10-item versions of WBIS/-M revealed better item characteristics, internal consistency, model fit indices for unidimensionality, and convergent and divergent validity than the original 11-item versions, with the 10-item version of the WBIS-M showing the best psychometric properties in the present sample.

Item analyses yielded a low percentage of missing item responses (<5%) for both the WBIS and WBIS-M as reported before [12]. Item-total correlations were medium- to large-sized, being smaller for the WBIS than for the WBIS-M. Plausibly, item 4, the only item focusing on the wish to lose weight, was homogenously answered with the maximum response option (highly) by prebariatric patients willing to undergo invasive weight loss treatment and showed corrected item-total correlations below the threshold (r_{it} <0.30). Notably, item 4 showed lower mean scores and sufficient corrected item-total correlations in previous studies among US American prebariatric adults [12] and adolescents [10]. Possibly, in the USA also patients with a lower wish to lose weight opt for bariatric surgery, as it requires no presurgical dietetic therapy and is more common than in Germany [38, 39]. In Germany, patients face high bureaucratic requirements [40] and must undergo an obligatory dietetic and behavioral weight loss therapy before surgery, which could

increase their wish to lose weight. Also different from US samples [10, 12], item 1 ("feeling competent") showed sufficient item-total correlations [41] in the present study and seems to be applicable for German prebariatric patients. This is consistent with lower competence-related teasing toward individuals with overweight and obesity in Germany compared to the USA [42].

As found previously [12], most items had a low kurtosis and negative skewness, indicating high item responses and agreement with WBI. The item difficulty indices ranged from very easy to very difficult and were easier than in a population-based sample [13]. This accords to the higher WBI and, especially counting for item 4, higher wish to lose weight among prebariatric samples compared to the general population. Notably, the items of the WBIS-M had more medium-sized difficulty indices than the WBIS, suggesting a better differentiation ability [27]. However, this may also be due to the more heterogeneous weight status of the WBIS-M than the WBIS subsample. Among all investigated versions, the new 10-item version of WBIS-M revealed the best item characteristics. Surprisingly, in the present study, the WBIS/-M mean scores after removing item 4 resembled those in US American prebariatric studies after removing item 1 [10, 12], but were smaller than those in a German prebariatric study after removing item 1 [11]. Most likely, the mean scores of the study by Hübner et al. [11] were positively skewed as the, in Germany, highly rated item 4 was retained, and the lowly rated item 1 was removed. Moreover, the WBIS yielded significantly higher mean scores than the WBIS-M as shown before in population-based samples [8, 13], which might reflect a lower identification of prebariatric patients with the weight-neutral language of the WBIS-M.

The unidimensionality of both versions of the WBIS was supported by two of five model fit indices, in accordance with previous evidence in prebariatric samples [12], whereas the unidimensionality of both versions of the WBIS-M was supported by three of five model fit indices. Among all tested versions, the 10-item version of WBIS-M showed the best model fit among prebariatric samples with a CFI close to the cut-off, being interpretable as supporting the model fit [43]. Nevertheless, the unidimensionality of the WBIS/-M stays questionable. Internal consistency was rated good for all investigated versions, as similarly shown for the WBIS in another German prebariatric sample [11]. Again, the WBIS-M exceeded the WBIS, and the 10-item version of the WBIS-M showed the best internal consistency.

Concerning convergent validity, the WBIS/-M showed significant, medium- to large-sized correlations with the measures for validation, confirming previous findings on correlations with depression, eating disorder psychopathology, and quality of life among prebariatric samples [2, 10], while extending evidence for general self-efficacy. Additionally, perceived weight-related childhood teasing was added as a construct of convergent validity and showed medium-sized correlations with the WBIS-M, but only small-sized correlations with the WBIS, which may be due to significantly different scores in the POTS assessing perceived weight-related childhood teasing. In addition, the WBIS had generally lower correlations with the measures for convergent validation than the WBIS-M. Expectedly, the WBIS/-M showed only small-sized correlations with attempts to dietary restriction [10]. The ANOVA revealed no significant main or interaction effects of sex, age, and weight status on the WBIS/-M mean scores, suggesting that WBI is not a matter of these sociodemographic and anthropometric variables as similarly suggested by previous studies [11, 12]. The fact that in a German population sample [13] significant variations of the WBIS were found by sex, weight status, and sex × age interaction could be explained by a higher variance regarding these indicators in the population and greater homogeneity in our prebariatric sample.

The strengths of the study included the multicenter design, the large sample which was representative for bariatric surgery in Germany regarding sex, age, and BMI [44], and the investigation of the WBIS-M, which had not been psychometrically evaluated among prebariatric samples before. The WBIS/-M was validated with internationally established self-report questionnaires. Moreover, response bias was limited by using objectively measured body weight and height and data collection was conducted independently from clinical procedures.

On the other hand, the study is limited by the significant differences of mean age and weight status distribution between the subsamples. Further, the subsample data were assessed at different time points with time differences up to 10 years and 5 months and thus may be influenced by unobserved time-depending covariates. Moreover, this study did not assess test-retest reliability, socioeconomic factors, which were shown to have an impact on WBI in the general population [13], and gender dimensions outside the binary system. Finally, the study results are only representative for a Western industrialized country and must be interpreted considering this sociocultural background.

Conclusion

The comparison of the psychometric properties of the WBIS and the WBIS-M and their 11-item and 10-item versions after removing item 4 among German prebariatric patients found the new 10-item version of the WBIS-M to be the most valid and reliable. The questionable unidimensionality of the WBIS/-M should be investigated in further studies and might be discarded in favor of another factor structure. Finally, the new 10-item version of WBIS-M should be further investigated and validated, considering socioeconomic factors to identify risk groups of WBI and including international comparisons to understand whether and why differences in psychometric properties occur. Moreover, longitudinal studies on WBI among prebariatric/bariatric samples are necessary, to evaluate prospective relations between WBI and psychopathology. Possibly, the new 10-item version of WBIS-M could help detect WBI during bariatric surgery, which could be approached psychotherapeutically, e.g., by cognitive-behavioral therapy [45], and in turn improve bariatric surgery outcomes [5]. Nevertheless, the valid and reliable assessment of WBI should help rise the public attention for weight bias and WBI and encourage politicians, health care professionals, and the public to diminish this health burden.

Acknowledgments

The authors thank Martina Beyrau, Ph.D. and Julius Becker, cand. med., for editing this manuscript. Moreover, the authors thank all participants for contributing their time and data to the Psychosocial Registry for Obesity Surgery (PRAC).

Statement of Ethics

The present study is embedded in the multicenter Psychosocial Registry for Obesity Surgery (PRAC), which is registered in the German Clinical Trials Register (DRKS00006749). The survey was approved by the Ethics Committee of the University of Leipzig (no. 356/11-ff) and carried out according to the Declaration of Helsinki. All recruited patients agreed to participate in the PRAC study by written informed consent.

Conflict of Interest Statement

All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf and declared that Professor Hilbert received support from the Federal Ministry of Education and Research during the conduct of the study; grants from the Federal Ministry of Education and Research, German Research Foundation, and Roland Ernst Foundation for Healthcare outside the submitted work; royalties for books on the treatment of eating disorders and obesity with Hogrefe and Kohlhammer; honoraria for workshops and lectures on eating disorders and obesity and their treatment; honoraria as editor of the *International Journal of Eating Disorders* and the journal *Psychotherapeut*; honoraria as a reviewer from Mercator Research Center Ruhr, Oxford University Press, and the German Society for Nutrition; and honoraria as a consultant for WeightWatchers, Zweites Deutsches Fernsehen, and Takeda. The other authors have no competing interests to report.

Funding Sources

The Integrated Research and Treatment Center AdiposityDiseases was supported by the German Federal Ministry of Education and Research (Grant No. 01EO1501). The funding source was not involved in designing and conducting the study; collecting, managing, analyzing, and interpreting the data; preparing, reviewing, or approving the manuscript; and deciding to submit the manuscript for publication.

Author Contributions

Conceptualization and methodology: S.S., C.H., J.E., R.S., and A.H.; data curation: T.M., J.S., F.S., S.K., A.D., C.H., J.E., R.S., and A.H.; formal analysis: S.S. and R.S.; funding acquisition and resources: A.H.; investigation: C.H., J.E., and R.S.; project administration and supervision: C.H., J.E., R.S., and A.H.; writing original draft: S.S.; writing reviews and editing: all authors.

Data Availability Statement

The authors confirm to share the study protocol, including potential amendments; the statistical code to generate the published results; and the data set from which the results were derived with the journal editor. The data that support the findings of this study are not publicly available due to ethical reasons but are available from the corresponding author upon reasonable request. Further inquiries can be directed to the corresponding author.

References

1 World Obesity Federation [Internet]. Weight stigma [cited 2023 April 5]. Available from: https://www.worldobesity.org/what-we-do/our-policy-priorities/weight-stigma/.

2 Durso LE, Latner JD. Understanding self-directed stigma: development of the weight bias internalization scale. Obesity. 2008;16(Suppl 2):80–S86.

3 Alimoradi Z, Golboni F, Griffiths MD, Broström A, Lin CY, Pakpour AH. Weight-related stigma and psychological distress: a systematic review and meta-analysis. Clin Nutr. 2020;39(7):2001–13.

4 Oltmanns JR, Rivera Rivera J, Cole J, Merchant A, Steiner JP. Personality psychopathology: longitudinal prediction of change in body mass index and weight post-bariatric surgery. Health Psychol. 2020;39(3):245–54.
5 Bennett BL, Lawson JL, Funaro MC, Ivezaj V. Examining weight bias before and/or after bariatric surgery: a systematic review. Obes Rev. 2022;23(11):e13500.
6 Cosentino C, Marchetti C, Monami M, Mannucci E, Cresci B. Efficacy and effects of bariatric surgery in the treatment of obesity: network meta-analysis of randomized controlled trials. Nutr Metab Cardiovasc Dis. 2021;31(10):2815–24.
7 O'Brien PE, Hindle A, Brennan L, Skinner S, Burton P, Smith A, et al. Long-term outcomes after bariatric surgery: a systematic review and meta-analysis of weight loss at 10 or more years for all bariatric procedures and a single-centre review of 20-year outcomes after adjustable gastric banding. Obes Surg. 2019;29(1):3–14.
8 Pearl RL, Puhl RM. Measuring internalized weight attitudes across body weight categories: validation of the modified weight bias internalization scale. Body Image. 2014;11(1):89–92.
9 Papadopoulos S, de la Piedad Garcia X, Brennan L. Evaluation of the psychometric properties of self-reported weight stigma measures: a systematic literature review. Obes Rev. 2021;22(8):e13267.
10 Roberto CA, Sysko R, Bush J, Pearl R, Puhl RM, Schvey NA, et al. Clinical correlates of the Weight Bias Internalization Scale in a sample of obese adolescents seeking bariatric surgery. Obesity. 2012;20(3):533–9.
11 Hübner C, Schmidt R, Selle J, Köhler H, Müller A, de Zwaan M, et al. Comparing self-report measures of internalized weight stigma: the weight self-stigma questionnaire versus the weight bias internalization scale. PLOS ONE. 2016;11(10):0165566.
12 Wagner AF, Butt M, Rigby A. Internalized weight bias in patients presenting for bariatric surgery. Eat Behav. 2020;39:101429.
13 Hilbert A, Baldofski S, Zenger M, Löwe B, Kersting A, Braehler E. Weight bias internalization scale: psychometric properties and population norms. PLoS One. 2014;9(1):e86303.
14 Thompson JK, Cattarin J, Fowler B, Fisher E. The Perception Of Teasing Scale (POTS): a revision and extension of the Physical Appearance Related Teasing Scale (PARTS). J Pers Assess. 1995;65(1):146–57.
15 Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord. 1994;16(4):363–70.
16 Hilbert A, Tuschen-Caffier B. Eating disorder examination. Tübingen: dgvt-Verlag; 2016.
17 Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D). Diagnostica. 2004;50(4):171–81.
18 Kolotkin RL, Crosby RD, Kosloski KD, Williams GR. Development of a brief measure to assess quality of life in obesity. Obes Res. 2001;9(2):102–11.
19 Mueller A, Holzapfel C, Hauner H, Crosby RD, Engel SG, Mühlhans B, et al. Psychometric evaluation of the German version of the impact of weight on quality of life-lite (IWQOL-Lite) questionnaire. Exp Clin Endocrinol Diabetes. 2011;119(2):69–74.
20 Schwarzer R, Jerusalem M. Causal and control believes. The general self-efficacy scale (GSE). In: Wright S, Johnston M, Weinman J, editors. Measures in health psychology: a user's portfolio. Anxiety, stress, and coping; 2010. Vol. 12. p. 329–45.
21 IBM Corp. IBM SPSS statistics for Windows. IBM Corp. Version 27.0. https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-27.
22 IBM Corp. SPSS AMOS, Version 29.0. IBM Corp. https://www.ibm.com/support/pages/downloading-ibm-spss-amos-29-0.
23 Wobbrock JO, Elkin LA, Higgins JJ, Findlater L, Gergle D, Matthew K. ARTool Version 2.1.2. Available from: https://depts.washington.edu/acelab/proj/art/index.html.
24 Mirzaei A, Carter SR, Patanwala AE, Schneider CR. Missing data in surveys: key concepts, approaches, and applications. Res Social Adm Pharm. 2022;18(2):2308–16.
25 Nunnally JC. Psychometric theory. New York City: Tata McGraw-Hill Education; 1994.
26 Lei M, Lomax RG. The effect of varying degrees of nonnormality in structural equation modeling. Struct Equ Model: A Multidiscip J. 2005;12(1):1–27.
27 Moosbrugger H, Kelava A. Testtheorie und Fragebogenkonstruktion. 2nd rev. Berlin: Springer; 2007.
28 Howell DC. Statistical methods for psychology. New York: Wadsworth; 1998.
29 Macho S, Andrés A, Saldaña C. Validation of the modified weight bias internalization scale in a Spanish adult population. Clin Obes. 2021;11(4):e12454.
30 Bentler PM, Wu EJ. EQS for Windows user's guide. Encino: Multivariate Software Inc; 1995.
31 Bollen KA, Stine RA. Bootstrapping goodness-of-fit measures in structural equation models. Socio Methods Res. 1992;21(2):205–29.
32 Roos JM. Confirmatory factor analysis. Los Angeles: SAGE; 2022.
33 Stevens JP. Applied multivariate statistics for the social sciences. 5th ed. New York: Routledge; 2009.
34 DeVellis RF, Thorpe CT. Scale development: theory and applications. Los Angeles: SAGE; 2022.
35 Cohen J. Quantitative methods in psychology: a power primer. Psychol Bull. 1992;112(1):155–9.
36 Wang Y, Rodríguez de Gil P, Chen YH, Kromrey JD, Kim ES, Pham T, et al. Comparing the performance of approaches for testing the homogeneity of variance assumption in one-factor ANOVA models. Educ Psychol Meas. 2017;77(2):305–29.
37 Wobbrock JO, Findlater L, Gergle D, Higgins JJ. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. Proceedings of the SIGCHI conference on human factors in computing systems. 2011; Vancouver, BC, New York: Association for Computing Machinery.
38 Clapp B, Ponce J, DeMaria E, Ghanem O, Hutter M, Kothari S, et al. American Society for Metabolic and Bariatric Surgery 2020 estimate of metabolic and bariatric procedures performed in the United States. Surg Obes Relat Dis. 2022;18(9):1134–40.
39 Deutsches Ärzteblatt [Internet]. Gutachten: Kassen behindern operative Adipositasbehandlung. [cited 2023 April 5]. Available from: https://www.aerzteblatt.de/nachrichten/135253/Gutachten-Kassen-behindern-operative-Adipositasbehandlungen/.
40 Luck-Sikorski C, Jung F, Dietrich A, Stroh C, Riedel-Heller SG. Perceived barriers in the decision for bariatric and metabolic surgery: results from a representative study in Germany. Obes Surg. 2019;29(12):3928–36.
41 Smith GT, McCarthy DM. Methodological considerations in the refinement of clinical assessment instruments. Psychol Assess. 1995;7(3):300–8.
42 Kim TJ, Makowski AC, von dem Knesebeck O. Obesity stigma in Germany and the United States – results of population surveys. PLOS ONE. 2019;14(8):e0221214.
43 Iacobucci D. Structural equations modeling: fit indices, sample size, and advanced topics. J Consum Psychol. 2010;20(1):90–8.
44 Thaher O, Driouch J, Hukauf M, Glatz T, Croner RS, Stroh C. Is development in bariatric surgery in Germany compatible with international standards? A review of 16 years of data. Updates Surg. 2022;74(5):1571–9.
45 Pearl RL, Puhl RM. Weight bias internalization and health: a systematic review. Obes Rev. 2018;19(8):1141–63.