



Examination of the Reliability and Validity of the Minnesota Multiphasic Personality Inventory-3 (MMPI-3) in a Preoperative Bariatric Surgery Sample

Ashleigh A. Pona^{1,4} · Ryan J. Marek² · Eva Panigrahi¹ · Yossef S. Ben-Porath³

Accepted: 9 August 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Presurgical psychological assessment of bariatric surgery candidates aims to identify psychosocial risk factors and provide treatment recommendations to facilitate optimal outcomes. Such assessment typically includes psychometric testing and a clinical interview. The Minnesota Multiphasic Personality Inventory (MMPI) has been commonly used as a broadband measure to assess a number of psychosocial domains in bariatric clinics. The newest version of the MMPI, the MMPI-3, was recently released. This study sought to (1) establish whether the MMPI-3 is comparable to the MMPI-2-RF in a sample of patients seeking bariatric surgery, (2) report reliability data for all MMPI-3 scale scores in this sample, and (3) explore associations between commonly used self-report symptom measures and substantive scales of the MMPI-3 to ascertain convergent and discriminant validity patterns. Six hundred and thirty-five presurgical patients completed the MMPI-3 in addition to the Patient Health Questionnaire-9 (PHQ-9), General Anxiety Disorder-7 (GAD-7), Alcohol Use Disorders Identification Test-Consumption (AUDIT-C), and Eating Disorder Examination-Questionnaire (EDE-Q). The majority (79.1%) of the sample was female, 65.5% was white, and 26.6% was Black. Scores on most of the MMPI-3 Emotional/Internalizing Dysfunction scales were meaningfully associated with the PHQ-9, GAD-7, and most EDE-Q subscales (except for Restraint). Meaningful discriminant patterns were observed as well. We conclude that the substantive scales of the MMPI-3 are reliable, comparable to their MMPI-2-RF counterparts, and evidence good convergent validity with extra-test measures assessing depression, anxiety, alcohol use, and eating disorder psychopathology in a preoperative bariatric sample.

Keywords MMPI-3 · Bariatric surgery · Assessment · Validity · Reliability

Although bariatric surgery is the most effective treatment for severe obesity in terms of long-term weight loss and reduction in medical comorbidities (Ahmed et al., 2018; Jakobsen et al., 2018; O'Brien et al., 2019), some patients experience suboptimal surgical outcomes (King et al., 2018, 2020).

The American Society for Metabolic and Bariatric Surgery (ASMBS) recommends a presurgical psychological assessment to identify psychosocial risk factors and provide recommendations to both the patient and multidisciplinary team that aim to facilitate the best outcome for the patient (Sogg et al., 2016). The ASMBS recommendations indicate that presurgical assessment should include psychometric testing in addition to the clinical interview, with the rationale that test data can aid in forming a more comprehensive clinical impression, provide information that may not be sufficiently covered or available within the time restrictions of the interview, and reveal information about the patient that may not have been disclosed during the interview (Sogg et al., 2016).

The Minnesota Multiphasic Personality Inventory (MMPI) instruments have been commonly used in bariatric surgery clinics as broadband measures to assess a range of relevant psychosocial domains (Bauchowitz et al., 2005; Walfish et al., 2007). The newest version of the test, the

✉ Ashleigh A. Pona
ashleigh.pona@osumc.edu

¹ Department of Psychiatry and Behavioral Health, Ohio State University Medical Center, Columbus, OH, USA

² Department of Primary Care and Clinical Medicine, Sam Houston State University College of Osteopathic Medicine, Conroe, TX, USA

³ Department of Psychological Sciences, Kent State University, Kent, OH, USA

⁴ Ohio State University Medical Center, 2050 Kenny Road, Morehouse Pavilion, Suite 1473, Columbus, OH 43221, USA

MMPI-3 (Ben-Porath & Tellegen, 2020a), was released in November 2020. Extensive research with the previous version, the MMPI-2-Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011), demonstrated good psychometric properties in bariatric samples, including predictive utility, reliability and validity, and replicable comparison group data (Marek et al., 2021; Tarescavage et al., 2013). The goals of the MMPI-3 revision were to collect new normative data representative of the 2020 census and enhance content coverage while building on the previous MMPI instruments' strong foundations (Ben-Porath & Tellegen, 2020a). The MMPI-3 consists of 335 items comprising 52 scales, which include 10 validity scales, 3 higher-order scales, 8 restructured clinical scales, 26 specific problem scales (within the domains of somatic/cognitive, internalizing, externalizing, and interpersonal), and 5 PSY-5 scales.

To date, the MMPI-2-RF has been a particularly useful psychometric tool in bariatric surgery psychological evaluations because it has bariatric norms. Currently, the MMPI-3 has a female only bariatric surgery candidate population as a standard comparison group; more data are needed to create a male bariatric surgery candidate standard comparison group. Thus, the current project which examines the validation of the MMPI-3 in both female and male bariatric surgery candidates is important to long-time users of the MMPI-2 and the MMPI-2-RF.

Ben-Porath and Tellegen (2020b) provide extensive data analyses demonstrating that the empirical correlates of MMPI-3 scale scores are comparable to those obtained with MMPI-2-RF versions of these scales. Although these analyses included a presurgical spine surgery candidate sample, findings were not reported for bariatric surgery candidates. Marek et al., (2021) have demonstrated clinical utility of the new MMPI-3 Eating Concerns-specific problem scale in assessing eating pathology in a postoperative bariatric surgery sample. Specifically, elevated scores on the Eating Concerns scale were associated with 6-year postoperative percent weight regain and higher scores on the Eating Disorder Examination-Questionnaire. However, the psychometric properties of the broader set of MMPI-3 scales within a pre-operative bariatric sample have yet to be examined.

The purpose of the current study was to establish whether the MMPI-3 is comparable to the MMPI-2-RF in a sample of patients seeking bariatric surgery. We also aimed to report reliability data for all MMPI-3 scale scores. Last, we sought to explore associations between commonly used self-report symptom measures and the MMPI-3 Emotional/Internalizing Dysfunction, Behavioral Externalizing Dysfunction, and a few additional Specific Problems Scales (such as Eating Concerns) for the purpose of examining convergent and discriminant validity patterns. The self-report symptom inventories utilized in this study included the Patient Health Questionnaire-9 (PHQ-9; Spitzer et al., 1999) to assess depression, General Anxiety Disorder-7 (GAD-7; Spitzer

et al., 2006) to assess anxiety, Alcohol Use Disorders Identification Test-Consumption (AUDIT-C; Bush et al., 1998) to assess alcohol use, and Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994) to assess eating disorder psychopathology.

We hypothesized that MMPI-3 scale scores would be similar across genders with the exception of the Behavioral/Externalizing Dysfunction scales for which men tend to score 5 T to 6 T score points higher than women (Ben-Porath & Tellegen, 2020b). We also hypothesized that MMPI-3 scale scores would be similar to those who were administered the MMPI-2-RF in a different sample (reported in Marek et al., 2014) of patients seeking bariatric surgery. It was also hypothesized that MMPI-3 scale scores would demonstrate reliability coefficients comparable to those reported in other samples (Ben-Porath & Tellegen, 2020b). Notably, that Cronbach's alpha will be ≥ 0.70 for larger scales and that mean inter-item correlations will be $\geq .15$ for shorter Specific Problems scales. Last, it was hypothesized that MMPI-3 scales assessing facets of depression, anxiety, alcohol/substance use, and disordered eating would demonstrate substantial correlations [at 0.30, a moderate correlation as defined by Cohen (1988)] with commonly used, brief self-report symptom measures assessing similar constructs (i.e., PHQ-Q, GAD-7, AUDIT-C, and EDE-Q, respectively).

Method

Participants

Participants included patients who were seeking bariatric surgery at a large academic medical center in the Midwest and, as part of standard medical care, met with a psychologist for an evaluation prior to having bariatric surgery. Participants were required to be 18 years or older and English-speaking.

The sample consisted of 649 patients. A total of 14 patients were removed from the study sample because they invalidated the MMPI-3 based on criteria outlined in the MMPI-3 Technical Manual (Ben-Porath & Tellegen, 2020b). Of those who produced valid protocols ($n = 635$), 502 (79.1%) were female and 133 (20.9%) were male. The mean age was 41.93 ($SD = 11.05$). Race breakdown was as follows: 65.5% were white, 26.6% were Black, and the rest (7.9%) identified as another race. In terms of highest level of education attained, 4.7% did not complete high school or a GED, 23.1% completed high school or had a GED, 31.7% completed some college, 10.9% had an associate's degree, 17.5% had a bachelor's degree, and 12.1% had a master's degree or higher. The majority of patients (90.9%) were presenting for an initial bariatric surgery and 9.1% were presenting for a revision. In terms of surgery preference, 44.9%

desired sleeve gastrectomy, 40.6% desired Roux-en-Y gastric bypass, 13.4% were undecided, and 1.1% indicated adjustable gastric banding (although this procedure is no longer offered at the study hospital). In our study, the two MMPI-3 scales assessing underreporting, Uncommon Virtues (L) and Adjustment Validity (K), were elevated in 22.2% and 37.8% of patients, respectively.

Measures

Minnesota Multiphasic Personality Inventory-3 (MMPI-3)

The MMPI-3 (Ben-Porath & Tellegen, 2020a) is a 335-item broadband self-report measure of psychopathology and personality normed based on projected 2020 census demographics. It takes approximately 25–35 min to complete via computer administration. The test is comprised of 52 scales, which include 10 validity scales and 42 substantive scales. The 42 substantive scales include 3 higher-order scales, 8 restructured clinical scales, 26 specific problem scales (within the domains of somatic/cognitive, internalizing, externalizing, and interpersonal), and 5 PSY-5 scales. The scale scores of the MMPI-3 yield good reliability and validity across samples (Ben-Porath & Tellegen, 2020b) and among presurgical psychological evaluation samples (Marek et al., 2022).

Patient Health Questionnaire-9 (PHQ-9)

The PHQ (Spitzer et al., 1999) is a self-report diagnostic tool for psychological disorders that assesses the areas of depression, anxiety, eating, alcohol, and somatoform symptoms and was derived from the Primary Care Evaluation of Mental Disorders (PRIME-MD), a diagnostic tool created by Pfizer following the publication of the DSM-III. The PHQ-9 (Kroenke & Spitzer, 2002) is the 9-item depression module from the PHQ and has commonly been used in medical populations, including bariatric surgery patients. The PHQ-9 has yielded evidence of good validity, reliability, and utility as a depression screening tool with bariatric surgery patients (Cassin et al., 2013; Marek et al., 2016). Cronbach's alpha for the PHQ-9 in the current sample was 0.86.

General Anxiety Disorder-7 (GAD-7)

The GAD-7 (Spitzer et al., 2006) is a 7-item self-report screener for anxiety. Although the GAD-7 has demonstrated good reliability in bariatric surgery samples (Atwood et al., 2021; de Zwaan et al., 2014; Koehler et al., 2020), there is fairly limited research on its psychometric properties in bariatric settings (Marek et al., 2016). Sockalingam et al. (2017) found that anxiety scores as assessed by the GAD-7 decreased at 1 and 2 years after bariatric surgery

as compared to presurgery scores. Cronbach's alpha for the GAD-7 in the current sample was 0.91.

Eating Disorder Examination-Questionnaire (EDE-Q)

The Eating Disorder Examination (EDE; Fairburn & Cooper, 1993) is a semi-structured clinical interview that measures psychopathology of eating disorders, specifically concerns with shape, weight, and binge eating behaviors (Guest, 2000). The EDE includes four subscales including Dietary Restraint, Eating Concern, Shape Concern, and Weight Concern. The EDE-Q (EDE-Q; Fairburn & Beglin, 1994) was derived from the EQE for clinical and research purposes, and includes 28-items that address eating disorder behaviors and cognitive symptoms. The EDE-Q generates the same four subscales (Dietary Restraint, Eating Concern, Weight Concern, and Shape Concern) as the EDE plus a global score. The global score measures the incidence and severity of eating disorder behaviors (Rand-Giovannetti et al., 2020). The EDE-Q is a measure that is easily accessible on the public domain and administered in a short period, less than 10 min (Marek et al., 2016). The EDE-Q is a widely used clinical instrument that has been validated in the bariatric surgery population. It demonstrates adequate to good psychometric properties including adequate reliability (Cronbach's alpha ranges from 0.72 to 0.95) and good concurrent/criterion-related validity among bariatric surgery samples (Elder et al., 2006; Kalarchian et al., 2000; Marek et al., 2016).

Alcohol Use Disorders Identification Test-Consumption (AUDIT-C)

The AUDIT-C (Bush et al., 1998) is a modified version of the Alcohol Use Disorders Identification Test (AUDIT; Bush et al., 1998). The AUDIT is a 10-item measure that was developed by the World Health Organization (WHO) in an effort to measure alcohol use and behaviors and alcohol-associated problems (King et al., 2012). Among bariatric patients, the instrument demonstrates good reliability and convergent validity coefficients (King et al., 2012; Marek et al., 2016; Mitchell et al., 2015). The AUDIT-C is a shorter validated instrument that measures alcohol consumption in the past 12 months using 3 items (Marek et al., 2016; Suzuki et al., 2012). The measure is a widely accepted screening tool that is often included as part of the presurgical process for patients pursuing bariatric surgery. AUDIT-C scores range from 0 to 12 with a score of ≥ 4 for men and of ≥ 3 for women indicating positive for hazardous alcohol use or an active alcohol use disorder (Bush et al., 1998; Ibrahim et al., 2019). Cronbach's alpha for the AUDIT-C in the current sample was 0.41, likely owing to limited variability and small item count for the scale. The mean inter-item correlation was 0.29 indicating good reliability.

Procedure

As part of standard clinical care at the study hospital, patients met with a psychologist for an evaluation prior to having bariatric surgery. These presurgical evaluations are risk assessments aimed to identify psychosocial factors that may diminish the outcome of bariatric surgery. The presurgical evaluations consisted of 1 h of psychological testing with questionnaires administered via computer, and then, immediately after, 1 h of face-to-face interview with the psychologist. Approximate administration times for the questionnaires are as follows: 25–35 min for MMPI-3 (Ben-Porath & Tellegen, 2020a), 5 min for PHQ-9, 5 min for GAD-7, 5–10 min for EDE-Q, and <5 min for AUDIT-C (Marek et al., 2016). Since the onset of the COVID-19 pandemic, and for the entirety of this study, both the testing and interview portions of the presurgical evaluations have been conducted remotely. As noted above, as part of testing, patients were administered the following self-report measures: MMPI-3, PHQ-9, GAD-7, EDE-Q, AUDIT-C, and a health psychology demographics questionnaire (age, sex, marital status, race/ethnicity, highest level of education, history of bariatric surgery, and type of bariatric surgery being pursued). The MMPI-3 was administered via Pearson Assessment's Q-global, which is a secure web-based scoring and reporting system. All other measures were administered via REDCap (Harris et al., 2009, 2019), which is secure web application for building and managing online surveys and databases. Remote administration of psychological testing was proctored as recommended (Corey & Ben-Porath, 2020). Patients were evaluated consecutively between November 2020 and May 2021. Use of data was approved by the medical center's Institutional Review Board.

Statistical Analyses

Due to the large amount of data in relation to the sample size, a conservative correction method was deemed appropriate for interpreting results. For each set of analyses where *p*-values were interpreted, a Bonferroni-corrected alpha was calculated when determining statistical significance.

Means and standard deviations for the MMPI-3 scale scores and external criteria broken down by gender were calculated and placed in Table 1. To identify whether there were meaningful gender differences, independent samples *t*-tests were calculated. Cohen's *d*s (0.20 = small effect, 0.50 = medium effect, 0.80 or greater = large effect) were also calculated for every independent samples *t*-test (Cohen, 1988). Because of numerous comparisons, a Bonferroni-corrected alpha was calculated (0.05/58) and differences were only deemed statistically significant if alpha was less than 0.0009.

Means and standard deviations from the comparable MMPI-2-RF scales reported in Marek et al. (2014) and the current sample's combined gender MMPI-3 scale scores are reported in Table 2. Independent samples *t*-tests were calculated to compare scale scores along with Cohen's *d* to establish effect sizes. A Bonferroni-corrected alpha was calculated (0.05/47) and differences were only deemed statistically significant if alpha was less than 0.0010.

Listed in Table 3 are reliability and standard error of measurement estimates. Internal consistency coefficients—including mean inter-item correlations and Cronbach's alphas—and standard error of measurements were calculated. Kuder-Richardson-20 calculations (Kuder & Richardson, 1937) were used to estimate Cronbach's alphas due to the dichotomous nature of the MMPI-3 items (True/False).

Pearson Product–Moment Correlations were then calculated among the external criteria and between the MMPI-3 scale scores and the external criteria (Table 4) to examine the convergent and discriminant validity. A Bonferroni-corrected alpha was calculated (0.05/37) and correlations were only deemed statistically significant if alpha was less than 0.0014.

Results

Descriptive Statistics

Presented in Table 1 are descriptive statistics for the MMPI-3 scale scores and external criteria broken down by gender. With regard to the MMPI-3 scale scores, men scored statistically significantly higher than women on Behavioral/Externalizing Dysfunction, Antisocial Behaviors, Juvenile Conduct Problems, Impulsivity, Aggression, Cynicism, and Disconstraint. Effect sizes for these differences were in the small to modest range. No other statistically significant gender differences on the MMPI-3 scale scores or external criteria were observed. Thus, data were combined for further analyses.

Table 2 provides scale score comparisons between a sample of patients seeking bariatric surgery who took the MMPI-2-RF (Marek et al., 2014) vs. the current sample that took the MMPI-3. Those who took the MMPI-2-RF scored trivially to modestly higher on the following scales: Infrequent Responses, Somatic Complaints, Neurological Complaints, and Aggressiveness. There was also substantial difference between Malaise scores such that those who took the MMPI-2-RF scored, on average, 12 T score points higher than those who took the MMPI-3.

Table 1 Means and standard deviations for MMPI-3 scales and other self-report criteria ($N=635$)

	% Elevated	Men		Women		Inferential Statistics		
		<i>(n = 133)</i>		<i>(n = 502)</i>		<i>t</i> (633)	<i>p</i> -value	Cohen's <i>d</i>
		M	SD	M	SD			
Validity								
Combined response inconsistency	N/A	48	9	47	9	.84	.401	.08
Variable response inconsistency	N/A	48	10	47	9	.46	.649	.04
True response inconsistency	N/A	52 ^F	9	51 ^F	9	.10	.922	.01
Infrequent responses	N/A	50	12	47	9	2.54	.011	.25
Infrequent psychopathology responses	N/A	48	10	45	7	3.11	.002	.30
Infrequent somatic responses	3.1	50	12	50	10	.47	.638	.05
Symptom validity	4.3	53	10	56	10	2.57	.010	.25
Response bias	3.8	53	11	51	10	1.62	.105	.16
Uncommon virtues	22.2	54	10	55	10	1.52	.129	.15
Adjustment validity	37.8	54	11	56	11	1.47	.142	.14
Higher-order								
Emotional/internalizing dysfunction	11.3	50	12	48	11	1.19	.235	.12
Thought dysfunction	4.4	47	10	46	9	1.45	.146	.14
Behavioral/externalizing dysfunction	.6	45	9	41	7	6.17	<.001*	.60
Restructured clinical								
Demoralization	10.9	49	11	48	10	.90	.366	.09
Somatic complaints	14.5	51	12	51	11	-.78	.436	-.08
Low positive emotions	11.5	51	12	49	10	1.93	.054	.19
Antisocial behavior	1.6	47	9	43	7	5.59	<.001*	.55
Ideas of Persecution	5.4	47	10	47	9	.45	.656	.04
Dysfunctional negative emotions	8.7	48	10	47	10	.95	.342	.09
Aberrant experiences	3.1	48	11	45	8	3.40	.001	.33
Hypomanic activation	2.8	46	9	43	8	3.45	.001	.34
Somatic-cognitive-specific problems scales								
Malaise	20.6	54	11	53	11	.41	.684	.04
Neurological complaints	11.5	50	12	49	10	.73	.463	.07
Eating concerns	2.8	51	10	48	9	3.18	.002	.31
Cognitive complaints	11.0	49	12	48	10	.37	.711	.04
Internalizing specific problems scales								
Suicidal/death ideation	11.5	47	8	46	5	1.77	.077	.17
Helplessness/hopelessness	5.5	47	10	44	8	2.59	.010	.25
Self-doubt	13.5	51	11	50	10	.57	.566	.06
Inefficacy	9.3	48	10	48	10	.27	.789	.03
Stress	9.8	49	11	49	10	.44	.659	.04
Worry	15.3	48	11	47	10	.80	.425	.08
Compulsivity	17.0	50	9	49	11	.95	.345	.09
Anxiety related experiences	10.9	48	12	50	11	2.22	.027	.22
Anger proneness	5.2	48	10	46	9	2.27	.024	.22
Behavior restricting fears	5.4	50	11	50	11	.16	.870	.02
Externalizing specific problems scales								
Family problems	6.8	47	10	45	9	2.05	.040	.20
Juvenile conduct problems	3.8	49	11	44	8	5.63	<.001*	.55
Substance abuse	.5	43	6	42	5	2.49	.013	.24
Impulsivity	6.0	48	11	45	8	3.52	<.001*	.34
Activation	6.0	47	9	46	9	.85	.394	.08
Aggression	1.3	46	9	44	6	3.56	<.001*	.35

Table 1 (continued)

	% Elevated	Men		Women		Inferential Statistics		
		<i>(n = 133)</i>		<i>(n = 502)</i>		<i>t(633)</i>	<i>p</i> -value	Cohen's <i>d</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Cynicism	5.5	47	10	43	9	5.25	<.001*	.51
Interpersonal specific problems scales								
Self-Importance	3.9	49	10	48	9	.87	.387	.08
Dominance	6.9	48	9	48	9	.77	.441	.08
Disaffiliativeness	8.5	48	10	48	9	.51	.613	.05
Social avoidance	17.2	52	11	52	11	.78	.435	.08
Shyness	10.2	50	11	49	10	1.08	.281	.11
Personality psychopathology-5								
Aggressiveness	1.1	47	9	45	7	2.98	.003	.29
Psychoticism	4.1	48	10	46	9	1.37	.171	.13
Disconstraint	.8	46	8	42	6	5.84	<.001*	.57
Negative emotionality/neuroticism	8.8	48	11	48	11	-.50	.619	.05
Introversiion/low positive emotionality	18.7	55	12	52	12	2.30	.022	.22
Patient health questionnaire—9	19.1	6.17	5.82	5.21	4.79	1.95	.051	.19
Generalized Anxiety Disorder—7	9.0	3.65	4.08	3.36	4.08	.72	.475	.07
EDE-Q—Restraint	N/A	2.52	1.44	2.38	1.30	1.10	.272	.11
EDE-Q—Eating concerns	N/A	.97	1.11	.84	1.05	1.20	.232	.12
EDE-Q—Shape concerns	N/A	3.16	1.56	3.02	1.47	.96	.339	.09
EDE-Q—Weight concerns	N/A	2.96	1.25	2.93	1.16	.29	.771	.03
EDE-Q—Global score	6.9	2.40	1.10	2.29	1.00	1.10	.271	.11
Alcohol use disorder identification test	6.0 (Men)/ 7.4 (Women)	1.12	1.51	.96	1.03	1.41	.158	.14

*Indicates *p*-value fell below the Bonferroni alpha correction of $\leq .0009$

n Sample Size, *M* Mean, *SD* Standard Deviation, *MMPI-3* Minnesota Multiphasic Personality Inventory – 3, *EDE-Q* Eating Disorder Examination-Questionnaire

Reliability Analyses

Regarding the internal consistency coefficients reported in Table 3, median Cronbach's alpha estimates for the Higher-Order Scales were 0.77. The median reliability estimate among the Restructured Clinical Scales was 0.80. The Specific Problems Scales yielded a median of 0.72. The median internal consistency estimate among the Personality-Psychopathology-5 Scales was 0.75.

Mean inter-item correlations for the Higher-Order Scales yielded a median of 0.13. Regarding the Restructured Clinical Scales, the mean inter-item correlation median was 0.20. Mean inter-item correlations for the Specific Problems Scales yielded a median of 0.27. Among the Personality-Psychopathology-5 Scales, mean inter-item correlations yielded a median of 0.14.

Standard Error of Measurements (SEMs) are expressed in *T*-scores in Table 3. Among the Higher-Order Scales, these SEMs yielded a median of 3.63. Among the Restructured Clinical Scales, the median SEM was 4.16. With regard to the Specific Problems Scales, SEMs yielded a median of

4.61. Among the Personality-Psychopathology-5 Scales, the median SEM was 4.37.

Pearson Product-Moment Correlations were conducted on the external criteria measures. The inter-correlations between the external criteria measures are reported in Supplemental Table A. A substantial correlation was observed between the PHQ-9 and GAD-7 ($r = 0.79$), implying that both measures are likely capturing a similar construct vs. discriminating between depression and anxiety. Likewise, large inter-correlations were observed within EDE-Q subscales.

Validity Analyses

Most of the Emotional/Internalizing Dysfunction scales of the MMPI-3 were meaningfully associated with the PHQ-9, GAD-7, and most EDE-Q subscales (except for Restraint), see Table 4. Some discriminant patterns can be observed despite high inter-correlations among the external criteria. For instance, Low Positive Emotions scores were more strongly associated with the PHQ-9 than with the GAD-7.

Table 2 Scale score comparability between valid MMPI-2-RF and MMPI-3 scale scores

	MMPI-2-RF**			MMPI-3			Inferential statistics		
	(n = 1,268)			(n = 635)			<i>t</i> (633)	<i>p</i> -value	Cohen's <i>d</i>
	% Elevated	M	SD	% Elevated	M	SD			
Validity									
Variable response inconsistency	N/A	48	9	N/A	47	9	.50	.617	.11
True response inconsistency	N/A	52F	9	N/A	51F	7	.53	.596	.12
Infrequent responses	N/A	53	12	N/A	48	10	2.26	.024	.45
Infrequent psychopathology responses	N/A	48	8	N/A	46	8	1.06	.289	.25
Infrequent somatic responses	4.9	52	12	3.1	50	11	.88	.377	.17
Symptom validity	3.2	58	11	4.3	55	10	1.39	.165	.29
Response bias	2.8	55	11	3.8	52	10	1.39	.165	.29
Uncommon virtues	36.4	57	11	22.2	55	10	.93	.355	.19
Adjustment validity	17.9	53	10	37.8	55	11	.93	.355	.19
Higher-order									
Emotional/internalizing dysfunction	11.7	50	11	11.3	49	11	.45	.651	.09
Thought dysfunction	4.8	48	9	4.4	46	9	1.00	.318	.22
Behavioral/externalizing dysfunction	2.8	45	9	.6	42	7	1.59	.112	.37
Restructured clinical									
Demoralization	12.2	52	10	10.9	48	11	1.85	.064	.38
Somatic complaints	25.6	57	12	14.5	51	11	2.65	.008	.52
Low positive emotions	15.4	52	11	11.5	49	11	1.36	.175	.27
Antisocial behavior	4.0	46	9	1.6	44	7	1.06	.289	.25
Ideas of persecution	11.5	51	10	5.4	47	9	1.95	.052	.42
Dysfunctional negative emotions	6.5	47	10	8.7	47	10	.00	> .999	.00
Aberrant experiences	4.6	47	9	3.1	46	9	.50	.617	.11
Hypomanic activation	2.3	43	9	2.8	44	8	.51	.607	.12
Somatic–cognitive-specific problems scales									
Malaise	45.8	65	12	20.6	53	11	5.30	< .001*	1.04
Neurological complaints	30.9	57	13	11.5	49	11	3.46	.001	.66
Cognitive complaints	11.1	51	11	11.0	48	11	1.36	.175	.27
Internalizing specific problems scales									
Suicidal/death ideation	14.0	49	9	11.5	46	6	1.64	.101	.39
Helplessness/hopelessness	5.1	47	9	5.5	45	8	1.03	.304	.23
Self-doubt	18.9	52	11	13.5	50	11	.90	.366	.18
Inefficacy	7.4	49	10	9.3	48	10	.47	.636	.10
Stress (compared to stress/worry on the MMPI-2-RF)	14.5	50	10	9.8	49	10	.47	.636	.10
Worry (Compared to Stress/Worry on the MMPI-2-RF)	14.5	50	10	15.3	47	10	1.42	.155	.30
Anxiety related experiences	8.7	49	10	10.9	50	11	.46	.644	.10
Anger proneness	9.6	48	10	5.2	46	9	.97	.331	.21
Behavior restricting fears	4.2	49	9	5.4	50	11	.47	.636	.10
Externalizing specific problems scales									
Family problems	6.5	48	10	6.8	45	9	1.46	.145	.32
Juvenile conduct problems	5.8	49	10	3.8	45	9	1.95	.052	.42
Substance abuse	.5	44	6	.5	42	5	1.28	.201	.36
Activation	4.6	44	9	6.0	46	9	1.00	.318	.22
Aggression	2.5	45	8	1.3	44	7	.55	.584	.13
Cynicism	10.1	48	10	5.5	44	9	1.95	.052	.42
Interpersonal specific problems scales									
Dominance (compared to interpersonal passivity on the MMPI-2-RF)	7.5	49	10	6.9	48	9	.49	.627	.11

Table 2 (continued)

	MMPI-2-RF**			MMPI-3			Inferential statistics		
	(n = 1,268)			(n = 635)			<i>t</i> (633)	<i>p</i> -value	Cohen's <i>d</i>
	% Elevated	M	SD	% Elevated	M	SD			
Disaffiliativeness	10.8	50	10	8.5	48	10	.95	.343	.20
Social avoidance	17.7	52	11	17.2	52	11	.00	> .999	.00
Shyness	8.9	47	9	10.2	49	10	.97	.331	.21
Personality psychopathology-5									
Aggressiveness	9.5	50	9	1.1	46	7	2.12	.034	.50
Psychoticism	4.3	47	9	4.1	47	9	.00	> .999	.00
Disconstraint	2.0	45	8	.8	43	7	1.09	.274	.27
Negative emotionality/neuroticism	8.9	49	10	8.8	48	11	.46	.644	.10
Introversion/low positive emotionality	15.1	53	11	18.7	53	12	.00	> .999	.00

*Indicates *p*-value fell below the Bonferroni alpha correction of $\leq .0010$

**Data were used from an independent sample for comparison purposes and were previously published in the following paper: Marek et al. (2014)

Scores on the Dysfunctional Negative Emotions scale and most of its facet scales, such as Worry and Negative Emotionality/Neuroticism, were more strongly associated with the GAD-7 than with the PHQ-9. The Eating Concerns scale on the MMPI-3 was most strongly associated the Eating Concerns subscale on the EDE-Q, though still meaningfully associated with the other EDE-Q subscales (except Restraint) and, as would be expected, not meaningfully correlated with the PHQ-9, GAD-7, or AUDIT-C—providing evidence of good discriminant validity. The MMPI-3 Substance Abuse scale was most strongly associated with the AUDIT-C and MMPI-3 scale scores evidenced good discriminant validity with the AUDIT-C.

Discussion

Use of the MMPI instruments is empirically supported in bariatric surgery settings (Marek et al., 2013, 2014; Tarescavage et al., 2013). This study adds to the existing literature by being the first to appraise the recently released MMPI-3 within a presurgical bariatric sample. Our findings indicate that the MMPI-3 is a psychometrically sound measure for presurgical bariatric psychological evaluations as discussed next.

MMPI-3 scale score differences between genders map onto most other samples reported in the MMPI-3 Technical Manual (Ben-Porath & Tellegen, 2020b). Both men and women produce comparable T score means and standard deviations except for some of the Behavioral/Externalizing Dysfunction scales where men tended to score higher than women. This also was a pattern observed on the MMPI-2-RF in bariatric seeking samples (Marek et al., 2013, 2014; Tarescavage et al., 2013). These differences likely

reflect actual differences rather than test bias; however, further studies using external criteria similar to Marek et al.'s (2014) with a bariatric surgery seeking sample are needed to directly address this question.

With regard to a comparison of MMPI-3 and MMPI-2-RF scales in bariatric surgery candidates, scores were similar on both the MMPI-2-RF and the MMPI-3 reflecting substantial cross-version comparability. This finding is consistent with data reported in Appendix E of the MMPI-3 Technical Manual (Ben-Porath & Tellegen, 2020b). Of note, MMPI-3 scale scores on most of the Somatic/Cognitive scales were modestly to substantially lower when compared to their MMPI-2-RF counterparts. Ben-Porath and Tellegen (2020b) report a comparison between the normative samples of the MMPI-2-RF (collected in the mid-1980s) and the MMPI-3 (collected in 2020), which demonstrates that there was a substantial increase in scores on the Somatic/Cognitive scales for the MMPI-3 normative sample. Thus, cross-version differences on the Somatic/Cognitive scales are accounted for by normative shifts, with MMPI-3 scale scores likely providing a more accurate reflection of somatic/cognitive functioning in medical samples compared to the MMPI-2-RF. Patients in the current sample produced MMPI-3 scores that are more in line with the MMPI-3 normative sample and this is a similar finding to those reported among patients seeking spine surgery (Marek et al., 2022).

Reliability data in the current sample are generally good. These findings are consistent with those reported in the MMPI-3 Technical Manual (Ben-Porath & Tellegen, 2020b) for the normative sample for most scales. There were some reliability estimates that are lower than conventional thresholds for adequate reliability (e.g., substance abuse). This is largely due to a restricted range of scores among patients seeking bariatric surgery, which attenuates reliability

Table 3 Minnesota Multiphasic Personality Inventory-3 Substantive Scale Internal Consistencies and Standard Errors of Measurement in a Bariatric Surgery Sample ($N=635$)

Scale	Mean inter-item correlation	Cronbach's alpha	SEM
Higher-order			
Emotional/internalizing dysfunction	.23	.89	3.63
Thought dysfunction	.13	.77	4.28
Behavioral/externalizing dysfunction	.12	.77	3.51
Restructured clinical			
Demoralization	.40	.91	3.17
Somatic complaints	.20	.84	4.50
Low positive emotions	.20	.78	5.00
Antisocial behavior	.15	.70	4.10
Ideas of persecution	.25	.81	4.03
Dysfunctional negative emotions	.26	.87	3.73
Aberrant experiences	.15	.72	4.69
Hypomanic activation	.15	.72	4.22
Somatic-cognitive-specific problems scales			
Malaise	.25	.71	6.03
Neurological complaints	.14	.61	6.63
Eating concerns	.20	.50	6.32
Cognitive complaints	.42	.89	3.52
Internalizing specific problems scales			
Suicidal/death ideation	.17	.51	3.97
Helplessness/hopelessness	.29	.70	4.61
Self-doubt	.47	.85	4.13
Inefficacy	.32	.80	4.33
Stress	.26	.67	5.67
Worry	.47	.86	3.83
Compulsivity	.31	.78	4.89
Anxiety related experiences	.30	.86	4.19
Anger proneness	.34	.85	3.48
Behavior restricting fears	.15	.53	7.25
Externalizing specific problems scales			
Family problems	.22	.72	4.78
Juvenile conduct problems	.27	.72	4.59
Substance abuse	.11	.52	3.61
Impulsivity	.30	.70	4.82
Activation	.19	.65	5.19
Aggression	.17	.48	5.10
Cynicism	.27	.83	3.84
Interpersonal specific problems scales			
Self-importance	.26	.77	4.27
Dominance	.18	.66	5.19
Disaffiliativeness	.39	.80	4.25
Social avoidance	.34	.82	4.61
Shyness	.34	.78	4.79
Personality psychopathology-5			
Aggressiveness	.11	.66	4.37
Psychoticism	.13	.69	4.97
Disconstraint	.14	.75	3.57
Negative emotionality/neuroticism	.36	.90	3.37
Introversiveness/low positive emotionality	.26	.83	4.79

SEM Standard Error of Measurement

Table 4 Pearson product-moment correlations between MMPI-3 scale scores and other self-report measures

	PHQ-9	GAD-7	EDE-Q Restraint	EDE-Q Eating concerns	EDE-Q Shape concerns	EDE-Q Weight concerns	EDE-Q Global score	AUDIT-C
EID	.68**	.70**	.13**	.46**	.51**	.54**	.50**	-.03
BXD	.28**	.31**	.03	.20**	.23**	.21**	.20**	.15**
RCd	.68**	.69**	.14**	.47**	.52**	.55**	.52**	.03
RC2	.50**	.45**	.05	.27**	.27**	.32**	.28**	-.08*
RC4	.24**	.25**	.01	.18**	.20**	.17**	.17**	.18**
RC7	.60**	.69**	.12**	.47**	.49**	.50**	.48**	.01
RC9	.24**	.31**	.01	.21**	.20**	.19**	.18**	.07*
EAT	.33**	.28**	.09*	.52**	.36**	.39**	.40**	.11**
SUI	.33**	.33**	.09*	.16**	.21**	.23**	.21**	-.01
HLP	.45**	.46**	.08*	.28**	.28**	.32**	.29**	-.04
SFD	.61**	.61**	.13**	.42**	.49**	.51**	.48**	-.04
NFC	.51**	.58**	.09**	.40**	.42**	.45**	.42**	.02
STR	.50**	.59**	.12**	.37**	.43**	.45**	.42**	.00
WRY	.55**	.67**	.12**	.41**	.49**	.51**	.47**	.07*
CMP	.36**	.45**	.14**	.27**	.32**	.31**	.32**	-.01
ARX	.53**	.65**	.10**	.38**	.42**	.45**	.41**	-.03
ANP	.46**	.50**	.07*	.34**	.31**	.33**	.32**	.04
BRF	.38**	.41**	.08*	.27**	.30**	.32**	.29**	.00
FML	.36**	.41**	.03	.28**	.24**	.24**	.24**	.07*
JCP	.22**	.22**	-.02	.14**	.16**	.13**	.13**	.10**
SUB	.09*	.08*	.03	.12**	.10**	.12**	.11**	.34**
IMP	.33**	.38**	.01	.27**	.25**	.27**	.24**	.07*
ACT	.19**	.26**	.01	.18**	.17**	.16**	.15**	.05
AGG	.27**	.35**	.10**	.21**	.23**	.25**	.24**	.11**
CYN	.34**	.39**	.11**	.28**	.32**	.29**	.31**	.02
AGGR	-.01	.02	.13**	.03	.05	.02	.07*	.07*
DISC	.24**	.26**	.02	.19**	.20**	.19**	.18**	.19**
NEGE	.57**	.70**	.11**	.42**	.47**	.50**	.46**	.03
INTR	.46**	.40**	.10**	.27**	.27**	.29**	.28**	-.13**

*indicates $p < .05$ **indicates p-value fell below the Bonferroni alpha correction of $\leq .0014$

MMPI-3 Minnesota Multiphasic Personality Inventory—3, *EID* Emotional/Internalizing Dysfunction, *BXD* Behavioral/Externalizing Dysfunction, *RCd* Demoralization, *RC2* Low Positive Emotions, *RC4* Antisocial Behavior, *RC7* Dysfunctional Negative Emotions, *RC9* Hypomanic Activation, *EAT* Eating Concerns, *SUI* Suicidal/Death Ideation, *HLP*(Helplessness/Hopelessness, *SFD* Self-Doubt, *NFC* Inefficacy, *STR* Stress, *WRY* Worry, *CMP* Compulsivity, *ARX* Anxiety-Related Experiences, *ANP* Anger Proneness, *BRF* Behavior-Restricting Fears, *FML* Family Problems, *JCP* Juvenile Conduct Problems, *SUB* Substance Abuse, *IMP* Impulsivity, *AGG* Aggression, *ACT* Activation, *CYN* Cynicism, *AGGR* Aggressiveness, *DISC* Disconstraint, *NEGE* Negative Emotionality/Neuroticism, *INTR* Introversion/Low Positive Emotionality, *PHQ-9* Patient Health Questionnaire—9, *GAD-7* Generalized Anxiety Disorder-7, *EDE-Q* Eating Disorder Examination—Questionnaire, *AUDIT-C* Alcohol Use Disorder Identification Test—Consumption

estimates. For some scales, such as the Eating Concerns scale, mean inter-item correlation coefficients are a better estimate of internal consistency. This is because Cronbach's alpha is impacted by the number of items on a scale. Nonetheless, most scales on the MMPI-3 yielded good reliability estimates in this sample. Standard errors of measurement correct for the attenuating effects of range restriction. Most standard error of measurements across the Higher-Order, Restructured Clinical, and PSY-5 Scales in this sample fall just at or below 5 T score points. This includes the Specific

Problems Scales that had lower reliability estimates, but some fall slightly above 6–7 T score points—a finding that is consistent with standard error of measurements reported for the MMPI-3 normative sample (Ben-Porath & Tellegen, 2020b).

Although the MMPI-3 Substance Abuse (SUB) scale correlated moderately with the AUDIT-C, the association was weaker than some of the other convergent correlations with other criteria. This is likely due to prevalence of alcohol use in the sample and the scope of both the AUDIT-C and

SUB scale on the MMPI-3. For instance, the AUDIT-C is intended to be a screener, not a full measure, of problematic alcohol use. The screener only contains three items, which limits the ability to assess the full range and severity of problematic alcohol use. Moreover, the AUDIT-C only assesses problematic alcohol use and not the wide range of substance abuse problems that the MMPI-3 SUB scale is able to capture. Finally, the SUB scale of the MMPI-3 is a face valid measure. Because approximately 20% of patients seeking bariatric surgery engage in an underreporting response style (Ambwani et al., 2013; Marek et al., 2015), scores on the SUB scale are likely range restricted as well. Nonetheless, the pattern of correlations indicate that SUB score can detect problematic alcohol and substance use in this population.

Regarding validity, there was evidence of convergent correlations between the Emotional/Internalizing Dysfunction scales and external criteria. For example, the MMPI-3 Emotional/Internalizing Dysfunction scales that assess Demoralization (and specific facets) correlated substantially with both the PHQ-9 and GAD-7. Low Positive Emotions correlated more strongly with the PHQ-9 compared to the GAD-7. Dysfunctional Negative Emotions (and its facets, notably Worry and Anxiety) correlated more highly with the GAD-7 vs. the PHQ-9. The Eating Concern scale on the MMPI-3 correlated highest with the Eating Concern scale on the EDE-Q.

An important consideration is that inter-correlations were high between the PHQ-9 and GAD-7 ($r=0.79$) and among EDE-Q subscales—findings that are not unique to this sample (Gideon et al., 2016; Rahman et al., 2022; Taube-Schiff et al., 2015; Teymoori et al., 2020). Both the PHQ-9 and GAD-7 scores and the EDE-Q Global score had the highest correlation with the MMPI-3 Demoralization (RCd) scale, indicating these measures are likely saturated with demoralization variance, limiting their ability to identify discriminating correlations between depression, anxiety, and core eating disorder constructs. This is likely due to the heterogeneity and symptom overlap of the diagnostic criteria and distress typically caused by eating disorder constructs (e.g., body image concerns). Indeed, Teymoori et al. (2020) also found a high correlation between the PHQ-9 and GAD-7 in their sample of patients post-traumatic brain injury. They stated that this may suggest a “unidimensional construct such that both instruments were part of a general common factor” (p. 12) because they were unable to independently explain the variance of the construct (Teymoori et al., 2020). They hypothesized that this may be due to there being a few similar items on both instruments, as well as the fact that depression and anxiety share some underlying aspects, including negative affect and negative bias in information processing (Teymoori et al., 2020). Interestingly, the Eating Concerns (EAT) scale on the MMPI-3 was not meaningfully associated with the EDE-Q Restraint subscale. This finding

is consistent with Marek et al.’s (2021, 2022) study which examined associations between the MMPI-3 EAT scale and the EDE-Q subscales in a postoperative bariatric sample. This likely reflects content overlap between the EAT scale and the other EDE-Q subscales of Eating Concern, Weight Concern, and Shape Concern. The EDE-Q Restraint subscale, on the other hand, overlaps in content with just one EAT scale item.

In terms of generalizability, the demographic makeup of our sample is similar to other bariatric surgery centers—that is, it was primarily comprised of women and the average age was between 40 and 45 (Welbourn et al., 2018). The majority (65.5%) of our sample was white, 26.6% was Black, and 7.9% was identified as another race. Of note, our results indicate a lack of gender differences on most MMPI-3 scales with the exception of some Behavioral/Externalizing Dysfunction (BXD) scales where men scored 4–5 T score points higher than women. These findings are consistent with similar patterns in the Technical Manual (Ben-Porath & Tellegen, 2020b) across other samples and likely reflect true gender differences. MMPI-3 scale scores in the current study are indeed similar to MMPI-2-RF scale scores in samples of patients seeking bariatric surgery.

The MMPI-3 assesses a broad number of psychosocial domains that are relevant and can be used to inform clinical impressions and recommendations in the preoperative bariatric surgery evaluation process. Our study demonstrates that the substantive scales of the MMPI-3 are reliable, comparable to their MMPI-2-RF counterparts, and have good convergent validity with extra-test measures assessing depression, anxiety, alcohol use, and eating disorder psychopathology. Additional research is needed to replicate our findings and continue to ascertain the psychometric qualities of the MMPI-3 in bariatric surgery settings. It is recommended that future research utilize different external criteria measures—such as data from the clinical interview and medical records—as well as outcome data to examine whether patterns of predictive validity evidenced with the MMPI-2-RF further generalize to the MMPI-3. In terms of clinical utility and deciding whether to add, continue to use, or eliminate the MMPI from bariatric psychological evaluation protocols, there are several points that clinicians may want to consider. First, the PHQ-9 and GAD-7 were derived from the Primary Care Evaluation of Mental Disorders (PRIME-MD) Patient Questionnaire which was developed for screening in primary care clinics to make referrals. These questionnaires, along with other brief symptom measures, typically do not assess beyond DSM criteria and no qualifications are required for administration. The MMPI utilizes construct-related assessment and assesses a broad range of psychological functioning with norms; however, qualifications and adequate training are required to use the MMPI.

The MMPI has also demonstrated incremental validity. For example, Martin-Fernandez et al. (2021) found that MMPI-2-RF scale scores accounted for an additional 3%–24% of the variability in postoperative eating behaviors and quality of life in bariatric surgery patients, above and beyond other preoperative variables including the EDE-Q, Binge Eating Scale, and interview portion of the psychological evaluation. Presurgical psychological evaluations are higher stake evaluations and, given the literature on the tendency for this population to present favorably (Ambwani et al., 2013; Marek, 2014), the validity scales can be helpful to assess for underreporting. Information on underreporting gathered from the validity scales can be integrated into the written report, communicated with members of the multidisciplinary team who also care for the patient, and discussed with the patient prior to or after the psychological evaluation. Discussion with the patient could help providers relay the importance of being open and honest during appointments in order for the team to make individualized, meaningful recommendations for the patient and ultimately increase the chances of optimal outcomes. Sharing this information with a patient can also help providers relay that they are interested in knowing if/when a patient is experiencing challenges before or after surgery so that they may be able to intervene with additional support/intervention.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10880-022-09908-2>.

Author Contributions AP contributed to conceptualization (equal), data curation (lead), formal analysis (supporting), investigation (lead), project administration (lead), and writing—original draft (lead). RM contributed to formal analysis (lead), methodology (lead), visualization (lead), and writing—original draft (lead). EP contributed to data curation (supporting), investigation (supporting), and writing—original draft (supporting). YBP contributed to conceptualization (equal), resources (lead), supervision (lead), and writing—review & editing (lead).

Funding Ryan Marek and Ashleigh Pona received Grant funding from the University of Minnesota Press to aid in collecting these data.

Data Availability Data are not available in repository but can be made available for replication purposes providing appropriate institutional agreements are met.

Code Availability Data analysis was conducted in SPSS.

Declarations

Conflict of interest Yossef Ben-Porath is a paid consultant to the MMPI-2-RF and MMPI-3 publisher, the University of Minnesota, and Distributor, Pearson. As co-author of the MMPI-2-RF and MMPI-3, he receives royalties on sales of these tests.

Ethical Approval The Ohio State Biomedical Sciences Institutional Review Board approved use of data and approved a Waiver of Consent Process and Full Waiver of HIPAA Research Authorization given the retrospective nature of the study and all data were collected as part of routine care.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

References

- Ahmed, B., King, W. C., Gourash, W., Belle, S. H., Hinerman, A., Pomp, A., Dakin, G., & Courcoulas, A. P. (2018). Long-term weight change and health outcomes for sleeve gastrectomy (SG) and matched Roux-en-Y gastric bypass (RYGB) participants in the Longitudinal Assessment of Bariatric Surgery (LABS) study. *Surgery, 164*(4), 774–783.
- Ambwani, S., Boeka, A. G., Brown, J. D., Byrne, T. K., Budak, A. R., Sarwer, D. B., Fabricatore, A. N., Morey, L. C., & O’Neil, P. M. (2013). Socially desirable responding by bariatric surgery candidates during psychological assessment. *Surgery for Obesity and Related Diseases, 9*(2), 300–305. <https://doi.org/10.1016/j.soard.2011.06.019>
- Atwood, M. E., Cassin, S. E., Rajaratnam, T., Hawa, R., & Sockalingam, S. (2021). The bariatric interprofessional psychosocial assessment of suitability scale predicts binge eating, quality of life and weight regain following bariatric surgery. *Clinical Obesity, 11*(1), e12421. <https://doi.org/10.1111/cob.12421>
- Bauchowitz, A. U., Gonder-Frederick, L. A., Olbrisch, M. E., Azarbad, L., Rye, M. Y., Woodson, M., ... & Schirmer, B. (2005). Psychosocial evaluation of bariatric surgery candidates: a survey of present practices. *Psychosomatic Medicine, 67*(5), 825–832.
- Ben-Porath, Y.S., & Tellegen, A. (2008/2011). *Minnesota Multiphasic Personality Inventory-Restructured Form: Manual for administration, scoring, and interpretation*. University of Minnesota Press
- Ben-Porath, Y. S., & Tellegen, A. (2020a). *The minnesota multiphasic personality inventory-3: Manual for administration, scoring, and interpretation*. University of Minnesota Press.
- Ben-Porath, Y. S., & Tellegen, A. (2020b). *The minnesota multiphasic personality inventory-3: Technical manual*. University of Minnesota Press.
- Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., Bradley, K. A., Ambulatory Care Quality Improvement Project (ACQUIP). (1998). The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking. *Archives of Internal Medicine, 158*(16), 1789–1795. <https://doi.org/10.1001/archinte.158.16.1789>
- Cassin, S., Sockalingam, S., Hawa, R., Wnuk, S., Royal, S., Taube-Schiff, M., & Okrainec, A. (2013). Psychometric properties of the Patient Health Questionnaire (PHQ-9) as a depression screening tool for bariatric surgery candidates. *Psychosomatics, 54*(4), 352–358. <https://doi.org/10.1016/j.psych.2012.08.010>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Psychology Press.
- Corey, D. M., & Ben-Porath, Y. S. (2020). Practical guidance on the use of the MMPI instruments in remote psychological testing. *Professional Psychology-Research and Practice, 51*(3), 199–204. <https://doi.org/10.1037/pro0000329>
- de Zwaan, M., Georgiadou, E., Stroh, C. E., Teufel, M., Koehler, H., Tengler, M., & Mueller, A. (2014). Body image and quality of life in patients with and without body contouring surgery following bariatric surgery: A comparison of pre- and post-surgery groups.

- Frontiers in Psychology*, 5, 1310. <https://doi.org/10.3389/fpsyg.2014.01310>
- Elder, K. A., Grilo, C. M., Masheb, R. M., Rothschild, B. S., Burke-Martindale, C. H., & Brody, M. L. (2006). Comparison of two self-report instruments for assessing binge eating in bariatric surgery candidates. *Behaviour Research and Therapy*, 44(4), 545–560. <https://doi.org/10.1016/j.brat.2005.04.003>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370.
- Fairburn, C. G., & Cooper, Z. (1993). The eating disorder examination. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (pp. 317–360). Guilford Press.
- Gideon, N., Hawkes, N., Mond, J., Saunders, R., Tchanturia, K., & Serpell, L. (2016). Development and psychometric validation of the EDE-QS, a 12 item short form of the Eating Disorder Examination Questionnaire (EDE-Q). *PLoS ONE*, 11(5), e0152744.
- Guest, T. (2000). Using the eating disorder examination in the assessment of bulimia and anorexia: Issues of reliability and validity. *Social Work in Health Care*, 31(4), 71–83. https://doi.org/10.1300/J010v31n04_05
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Ibrahim, N., Alameddine, M., Brennan, J., Sessine, M., Holliday, C., & Ghaferi, A. A. (2019). New onset alcohol use disorder following bariatric surgery. *Surgical Endoscopy and Other Interventional Techniques*, 33(8), 2521–2530. <https://doi.org/10.1007/s00464-018-6545-x>
- Jakobsen, G. S., Småstuen, M. C., Sandbu, R., Nordstrand, N., Hofsø, D., Lindberg, M., Hertel, J. K., & Hjelmæsæth, J. (2018). Association of bariatric surgery vs medical obesity treatment with long-term medical complications and obesity-related comorbidities. *JAMA*, 319(3), 291–301.
- Kalarchian, M. A., Wilson, G. T., Brolin, R. E., & Bradley, L. (2000). Assessment of eating disorders in bariatric surgery candidates: Self-report questionnaire versus interview. *International Journal of Eating Disorders*, 28(4), 465–469. [https://doi.org/10.1002/1098-108x\(200012\)28:4%3c465::aid-eat17%3e3.3.co;2-u](https://doi.org/10.1002/1098-108x(200012)28:4%3c465::aid-eat17%3e3.3.co;2-u)
- King, W. C., Chen, J. Y., Mitchell, J. E., Kalarchian, M. A., Steffen, K. J., Engel, S. G., Courcoulas, A. P., Pories, W. J., & Yanovski, S. Z. (2012). Prevalence of alcohol use disorders before and after bariatric surgery. *Jama-Journal of the American Medical Association*, 307(23), 2516–2525. <https://doi.org/10.1001/jama.2012.6147>
- King, W. C., Hinerman, A. S., Belle, S. H., Wahed, A. S., & Courcoulas, A. P. (2018). Comparison of the performance of common measures of weight regain after bariatric surgery for association with clinical outcomes. *Jama-Journal of the American Medical Association*, 320(15), 1560–1569. <https://doi.org/10.1001/jama.2018.14433>
- King, W. C., Hinerman, A. S., & Courcoulas, A. P. (2020). Weight regain after bariatric surgery: A systematic literature review and comparison across studies using a large reference sample. *Surgery for Obesity and Related Diseases*, 16(8), 1133–1144. <https://doi.org/10.1016/j.soard.2020.03.034>
- Koehler, H., Dorozhkina, R., Gruner-Labitzke, K., & de Zwaan, M. (2020). Specific health knowledge and health literacy of patients before and after bariatric surgery: A cross-sectional study. *Obesity Facts*, 13(2), 166–178. <https://doi.org/10.1159/000505837>
- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9), 509–515. <https://doi.org/10.3928/0048-5713-20020901-06>
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160.
- Marek, R. J. (2014). Assessing psychosocial functioning of bariatric surgery candidates with the Minnesota multiphasic personality inventory-2 restructured form (MMPI-2-RF) [Master's thesis, Kent State University]. Ohio Library and Information Network. Retrieved August 7, 2022 from https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=kent1374680793&disposition=attachment
- Marek, R. J., Ben-Porath, Y. S., Ashton, K., & Heinberg, L. J. (2014). Minnesota multiphasic personality inventory-2 restructured form (MMPI-2-RF) scale score differences in bariatric surgery candidates diagnosed with binge eating disorder versus BMI-matched controls. *International Journal of Eating Disorders*, 47(3), 315–319.
- Marek, R. J., Ben-Porath, Y. S., Windover, A., Tarescavage, A. M., Merrell, J., Ashton, K., Lavery, M., & Heinberg, L. J. (2013). Assessing psychosocial functioning of bariatric surgery candidates with the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF). *Obesity Surgery*, 23(11), 1864–1873.
- Marek, R. J., Block, A. R., & Ben-Porath, Y. S. (2022). Reliability and validity of Minnesota Multiphasic Personality Inventory-3 (MMPI-3) scale scores among patients seeking spine surgery. *Psychological Assessment*. <https://doi.org/10.1037/pas0001096>
- Marek, R. J., Heinberg, L. J., Lavery, M., Rish, J. M., & Ashton, K. (2016). A review of psychological assessment instruments for use in bariatric surgery evaluations. *Psychological Assessment*, 28(9), 1142–1157. <https://doi.org/10.1037/pas0000286>
- Marek, R. J., Martin-Fernandez, K., Heinberg, L. J., & Ben-Porath, Y. S. (2021). An investigation of the eating concerns scale of the Minnesota Multiphasic Personality Inventory-3 (MMPI-3) in a postoperative bariatric surgery sample. *Obesity Surgery*, 31(5), 2335–2338. <https://doi.org/10.1007/s11695-020-05113-y>
- Marek, R. J., Tarescavage, A. M., Ben-Porath, Y. S., Ashton, K., Merrell, J., & Heinberg, L. J. (2015). Using presurgical psychological testing to predict 1-year appointment adherence and weight loss in bariatric surgery patients: Predictive validity and methodological considerations. *Surgery for Obesity and Related Diseases*, 11(5), 1171–1181. <https://doi.org/10.1016/j.soard.2015.03.020>
- Martin-Fernandez, K. W., Marek, R. J., Heinberg, L. J., & Ben-Porath, Y. S. (2021). Six-year bariatric surgery outcomes: the predictive and incremental validity of presurgical psychological testing. *Surgery for Obesity and Related Diseases*, 17(5), 1008–1016.
- Mitchell, J. E., Steffen, K., Engel, S., King, W. C., Chen, J. Y., Winters, K., Sogg, S., Sondag, C., Kalarchian, M., & Elder, K. (2015). Addictive disorders after Roux-en-Y gastric bypass. *Surgery for Obesity and Related Diseases*, 11(4), 897–905. <https://doi.org/10.1016/j.soard.2014.10.026>
- O'Brien, P. E., Hindle, A., Brennan, L., Skinner, S., Burton, P., Smith, A., Crosthwaite, G., & Brown, W. (2019). 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. *Obesity Surgery*, 29(1), 3–14.
- Rahman, M. A., Dhira, T. A., Sarker, A. R., & Mehareen, J. (2022). Validity and reliability of the Patient Health Questionnaire scale (PHQ-9) among university students of Bangladesh. *PLoS ONE*, 17(6), e0269634.
- Rand-Giovannetti, D., Cicero, D. C., Mond, J. M., & Latner, J. D. (2020). Psychometric properties of the Eating Disorder

- Examination-Questionnaire (EDE-Q): A confirmatory factor analysis and assessment of measurement invariance by sex. *Assessment*, 27(1), 164–177. <https://doi.org/10.1177/1073191117738046>
- Sockalingam, S., Hawa, R., Wnuk, S., Santiago, V., Kowgier, M., Jackson, T., Okrainec, A., & Cassin, S. (2017). Psychosocial predictors of quality of life and weight loss two years after bariatric surgery: Results from the Toronto Bari-PSYCH study. *General Hospital Psychiatry*, 47, 7–13. <https://doi.org/10.1016/j.genhosppsych.2017.04.005>
- Sogg, S., Lauretti, J., & West-Smith, L. (2016). Recommendations for the presurgical psychosocial evaluation of bariatric surgery patients. *Surgery for Obesity and Related Diseases*, 12(4), 731–749. <https://doi.org/10.1016/j.soard.2016.02.008>
- Spitzer, R. L., Kroenke, K., Williams, J. B., Patient Health Questionnaire. (1999). Validation and utility of a self-report version of PRIME-MD - The PHQ primary care study. *Jama-Journal of the American Medical Association*, 282(18), 1737–1744. <https://doi.org/10.1001/jama.282.18.1737>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Suzuki, J., Haimovici, F., & Chang, G. (2012). Alcohol use disorders after bariatric surgery. *Obesity Surgery*, 22(2), 201–207.
- Tarescavage, A. M., Wygant, D. B., Boutacoff, L. I., & Ben-Porath, Y. S. (2013). Reliability, Validity, and Utility of the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) in Assessments of Bariatric Surgery Candidates. *Psychological Assessment*, 25(4), 1179–1194. <https://doi.org/10.1037/a0033694>
- Taube-Schiff, M., Van Exan, J., Tanaka, R., Wnuk, S., Hawa, R., & Sockalingam, S. (2015). Attachment style and emotional eating in bariatric surgery candidates: The mediating role of difficulties in emotion regulation. *Eating Behaviors*, 18, 36–40.
- Teymoori, A., Gorbunova, A., Haghish, F. E., Real, R., Zeldovich, M., Wu, Y. J., Polinder, S., Asendorf, T., Menon, D., & Steinbüchel, N. V. (2020). Factorial structure and validity of depression (PHQ-9) and anxiety (GAD-7) scales after traumatic brain injury. *Journal of Clinical Medicine*, 9(3), 873.
- Walfish, S., Vance, D., & Fabricatore, A. N. (2007). Psychological evaluation of bariatric surgery applicants: Procedures and reasons for delay or denial of surgery. *Obesity Surgery*, 17(12), 1578–1583.
- Welbourn, R., Pournaras, D. J., Dixon, J., Higa, K., Kinsman, R., Ottosson, J., Ramos, A., van Wagenveld, B., Walton, P., Weiner, R., & Zundel, N. (2018). Bariatric surgery worldwide: Baseline demographic description and one-year outcomes from the Second IFSO Global Registry Report 2013–2015. *Obesity Surgery*, 28(2), 313–322. <https://doi.org/10.1007/s11695-017-2845-9>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.