NIH Public Access
Author Manuscript
Obesity (Silver Spring). Author manuscript; available in PMC 2011 December 1.

Published in final edited form as:
Obesity (Silver Spring). 2011 December ; 19(12): 2336–2344. doi:10.1038/oby.2011.96.

Psychometric Properties and Construct Validity of the Weight-Related Eating Questionnaire in a Diverse Population

Susan M. Schembre[1] and **Karly S. Geller**[1]
[1]University of Hawai'i Cancer Center, Prevention and Control Program, Honolulu, Hawaii, USA

Abstract

This study evaluates the 16-item, four-factor Weight-Related Eating Questionnaire (WREQ), which assesses theory-based aspects of eating behavior, across diverse, nonclinical subgroups. A total of 621 men and women aged 18–81 years (34.3 ± 16.4) with a mean BMI of 25.7 ± 6.1 kg/m^2 (range 15.5–74.1 kg/m^2) were recruited from general education classes at the University of Hawai'i, Manoa and an online survey panel of Hawai'i residents to complete a web-based survey. Participants were predominantly white (23%), Asian/Asian-mix (42%), or Native Hawaiian/Pacific Islander (18%). The WREQ's factor structure was successfully replicated by confirmatory factor analysis (CFA) for the entire sample and by weight status, gender, age, and race with strong internal consistency. Four-week test–retest reliability (n = 31) for the subscales was excellent with interclass correlations of 0.849–0.932. Tests of population invariance confirmed the generalizability of the WREQ across all subgroups having provided no evidence that the factor structure, factor loadings, or indicator intercepts varied significantly between the groups. Multivariate regression analyses showed that emotional eating was independently associated with BMI (β = 0.272, P < 0.001) as well as moderate- and long-term weight change rates (weight gain) in young adults (β = 0.152, P = 0.042) and adults (β = 0.217, P = 0.001). Compensatory restraint was negatively associated with weight gain in adults (β = −0.133, P = 0.039). Routine restraint and emotional eating were highest among dieters. All associations remained significant after accounting for gender, age, and race. The hypothesized WREQ measurement model demonstrated very good construct validity, confirming the unbiased generalizability of the WREQ measure across sex, age, race, and BMI subgroups, and excellent criterion-related validity with respect to current BMI, weight change, and weight control status.

Introduction

The dramatic increase in the prevalence of obesity has been well documented (1,2). Obesity researchers from an array of health-related fields have since concentrated their efforts on determining which behavioral, genetic, social, and/or environmental factors have the greatest impact on weight gain. Perhaps one of the more promising fields of research is the investigation of cognitive mechanisms that mediate the link between behavior and obesity. Of particular interest is the observation that the act of initiating an eating episode is not only controlled by physiological demands, but it is also influenced by how we respond to cues from the environment and manage negative affect. This observation is particularly salient today as it is estimated adults make over 200 food-related decisions on a daily basis (3). Research shows many of these decisions are only loosely connected with a physiological

Correspondence: Susan M. Schembre (sschembre@crch.hawaii.edu).

Disclosure
The authors declared no conflict of interest.

need to replenish energy stores (physiological hunger) (4). Mechanisms that promote eating in the absence of physiological hunger are believed to be the link between energy intake and obesity (5). The identification and assessment of nonphysiologically driven cognitive and psychological mechanisms to eat provides a more comprehensive understanding of the aspects of eating behavior that promote obesity. Researchers have developed and tested multiple hypotheses based on laboratory observations and clinical research in an effort to explain why some individuals are more susceptible to becoming obese. From these hypotheses, theory-based psychometric instruments have been designed to help identify individuals at risk for becoming obese.

Three eating behavior theories have emerged since 1950 that have led to a greater understanding of the underlying cognitive and psychological aspects of eating behavior that promote obesity. The psychosomatic theory (6), the theory of externality (7), and the theory of restraint (8) have provided the foundation on which weight-related eating behavior questionnaires have been developed. From these theories multiple questionnaires have been developed to assess the extent to which individuals' eating is motivated by cognitive and psychological factors including, but not limited to, Herman and Polivy's Restraint Scale (9), the 51-item Three Factor Eating Questionnaire (TFEQ) (10), the 33-item Dutch Eating Behavior Questionnaire (11), the 18-item and 21-item Revised TFEQs (12–14). Additionally, Lowe and colleagues recently developed the Power of Food Scale (15,16) that assesses the psychological impact of living in food-abundant environments. All of these psychometric instruments continue to be used in eating behavior research today to assess cognitive and psychological aspects of eating behavior. However, researchers should take care in selecting the most appropriate questionnaire for their specific study population as previously published validation studies have raised concerns about their use outside of the population in which they were developed. Some of the more limiting problems have been the failure to replicate the factor structures of the Restraint Scale (17) and the TFEQ (14,18) and the inability to validate scales of dietary restraint against objective measures of caloric intake (19–21).

In response to evidence of weak validity, researchers have sought to improve on the psychometric properties of these instruments. Revisions of the TFEQ have proven to be most successful both with respect to the validation of the instrument and their consistency with the underlying theoretical constructs on which the questionnaires were based (12,13). Subscales on both questionnaires have been shown to be associated with body weight status and dietary intakes (12,13,22,23). However, limitations still remain, particularly with regard to subject burden as a result of questionnaire length and the inclusion of questionnaire items that reference a food (e.g., "sizzling steak") that may be unappealing to a specific subgroup (e.g., vegetarians). Furthermore, none of these questionnaires have been tested for factorial invariance across ethnically diverse subgroups.

Most recently, the 16-item Weight-Related Eating Questionnaire (WREQ) (24) was developed to address shortcomings in each of the above referenced psychometric instruments. The WREQ is most similar to the Dutch Eating Behavior Questionnaire and the TFEQ-R in that it assesses susceptibility to external cues to eat (external eating) and emotional eating. However, the WREQ captures an additional construct of dietary restraint. Specifically, the WREQ distinguishes between restrictive eating patterns related to restrictive dieting (routine restraint) and a more flexible approach to weight control characterized by intentional caloric restriction following periods of perceived overconsumption (compensatory restraint). Evidence for subscales of dietary restraint has been previously addressed in the literature (17,25) and are, by definition, consistent with the subscales measured by the WREQ. Assessing two constructs of dietary restraint is of particular importance given that the direction of their associations with patterns of dietary intake and BMI have been demonstrated to oppose one another in some populations (25,26). The WREQ is the first questionnaire to assess all four constructs on one scale.

The observed factor structure of the WREQ was derived from a large sample of undergraduate university students in the northeastern United States, which resulted in a four-factor structure that a confirmatory factor analysis (CFA) further replicated in a corresponding sample of university students (24). Prior to using the WREQ among individuals from a different subgroup of the population, the psychometric properties and generalizability of the WREQ factor structure across subgroups requires examination. Study outcomes based on psychometric instruments that do not demonstrate factorial validity and invariance may be biased and result in a detection of shifts in measurement rather than true differences in the assessed constructs. The purpose of the current study is to evaluate the psychometric properties (factorial validity/reliability and test–retest reliability), population factorial invariance by gender, age, race, and weight status subgroups, and criterion-related validity of the WREQ in a multiethnic sample of (nonclinical) adults and young adults of various weight-status living in Hawai'i. We hypothesized there will be no evidence to conclude that the factor structure of the WREQ are different between groups and that the subscales can differentiate between those who are weight gaining (external and/or emotional eating scales) or are attempting to control their weight (compensatory and/or routine restraint).

Methods and Procedures

Participants

Two distinct and ethnically diverse, nonclinical study populations were recruited to complete an online survey. The first sample of participants included year-round Hawai'i residents (aged 25–81 years) recruited using The Hawai'i Panel online survey that is supported by Ward Research (Honolulu, HI). The Hawai'i Panel is comprised of part- and fulltime residents of all the islands of Hawai'i who have volunteered to receive electronic invitations to complete online surveys. Those members who were permanent Hawai'i residents for at least 6 months and over 24 years old were sent invitations to complete the WREQ survey. Consenting participants completed a series of screening questions to establish eligibility before being asked to complete the remainder of the survey. Individuals were deemed ineligible if they reported having a chronic health condition that required special dietary restrictions (e.g., Celiac disease, type 1 or 2 diabetes), had a self-reported history of a prior or current eating disorder diagnosis, or were currently pregnant or lactating. Eligible participants were compensated with PayPal cash or store gift cards equivalent to $10. The second sample of participants included young adults (18–24 years) who were recruited from general education courses at a local university. All consenting students, regardless of eligibility status, completed the WREQ survey and received class credit as compensation. Those students not meeting the eligibility criteria were later identified and excluded from all data analyses. All materials and protocols were approved by the institutional review board of the University of Hawai'i at Manoa.

Measures

All survey data was collected and organized via SurveyMonkey at http://www.surveymonkey.com/ (Portland, OR), an online survey software and questionnaire tool. In addition to WREQ items, the survey assessed information relevant to eligibility criteria, participant demographic characteristics, and current weight and height and weight history.

WREQ—The WREQ (24) was developed to measure cognitive and emotional aspects of eating as they relate to body weight. As depicted in Figure 1, the questionnaire consists of 16 items indicating four, correlated, but independent constructs of eating behavior: compensatory restraint (three items), routine restraint (three items), susceptibility to external cues (five items), and emotional eating (five items). The compensatory restraint construct reflects the intentional restriction of energy intake following an episode of overeating. The routine restraint construct reflects perceived routine restriction of energy intake to control weight. The susceptibility to external cues construct refers to eating in response to external orosensory cues without regard for internal signals of hunger or satiety. The emotional eating construct represents eating in response to negative emotions. Participants were asked to rate how well each statement described them on a five-point scale with higher ratings reflecting stronger agreement with the statement.

Participant characteristics—Demographic variables included self-reported gender, age, and race. Race data was collected in such a manner that participants could select multiple (provided) race categories and/or write in a category that was not listed as a response option. Eight categories were provided as response options for race based on the race/ethnic composition of Hawai'i and previously developed surveys (27): Black/African American, white, Chinese, Filipino, Native Hawaiian, Japanese, Korean, and Mexican or other Hispanic. The coded race variable used in this study reflects four main groups: white; Asian (single-race)/Asian mix, including single-race and mixed-race Filipino (Hispanic and non-Hispanic), Chinese, Japanese, Korean, other Asian (e.g., Thai and Vietnamese); Native Hawaiian/Pacific Islanders, including all individuals who selected Native Hawaiian or identified themselves as Pacific Islander (in the "other" category); and all "other mixed races", which included all other individuals.

Weight status—Weight and height were collected with the following questions to calculate current BMI: (i) "To the best of your knowledge, what is your current weight in pounds?" and (ii) "To the best of your knowledge, what is your height in feet and inches?" Current BMI was later converted to kg/m^2.

Weight control status was assessed by asking participants to identify themselves as one of the following response options: (i) I am not currently concerned with my weight; (ii) I am trying to avoid weight gain; (iii) I am dieting to lose weight; (iv) I am trying to gain weight.

Weight change rate was calculated for young adults and adults retrospectively as a proxy for moderate- and long-term energy balance, respectively. Young adults were asked what their weight was at the beginning of the previous Fall semester (5 months prior to current study). Adults were asked their weight at 21 years of age. Weight change rate was calculated for young adults as weight change (current−previous weight) divided by 5 (months) and is reported as kg/month. For adults, weight change rate was calculated as weight change since 21 years of age divided by the difference between their current reported age and age 21 years and is reported as kg/year. Accordingly, positive values reflect weight gain.

Data analysis

Analyses evaluated the validity and reliability of the WREQ, including descriptive statistics, tests of the psychometric properties (factor validity, internal consistency, and test–retest reliability), population factorial invariance testing, and criterion-related validation. Based on previously published data (24), the four-factor WREQ structure described earlier (see Figure 1) was examined further using CFA and nested model comparisons. The factorial validity of the WREQ model was first examined via model fit statistics, followed by nested model comparisons that tested factorial invariance across the selected subgroups. Prior to tests of model fit, subgroups were created for specific demographic variables. Given the influence of sample size on the χ^2 statistic and several related statistics (e.g., standard errors, parameter estimates), unbalanced subgroups were corrected by randomly selecting cases from the larger group. Binary subgroups were created for gender (males, females; n = 242), age (young adults, adults; n = 266), BMI (underweight/normal weight, overweight/obese; n = 292), and the two largest race groups (Asians/Asian-mix, whites; n = 144). Missing data (<1%) were handled with listwise deletion.

Factor structure and internal consistency

The validity factor structure of the WREQ was examined using CFA in Mplus Version 5.21 (28). Multiple estimates were used to judge appropriate model fit, including χ^2, the comparative fit index (CFI) (29), the standardized root mean square residual (SRMR), the root mean square error of approximation with CFit statistic, and the Tucker–Lewis index (30,31). An acceptable model fit was supported by a CFI value ≥0.90 (29), SRMR value ≤0.08, root mean square error of approximation value ≤0.05 and nonsignificant CFit statistic (P > 0.05) (32), and Tucker–Lewis index ≥0.95 (33). Estimates of factor loadings, intercepts, variances, residual variances, and z-scores (>1.96) were inspected for direction of association and magnitude. Factor determinacy coefficients were estimated as a measure of internal consistency, with values ranging from 0 to 1 and larger values indicating better measurement of the factor by the observed items. A factor determinacy coefficient of ≥0.80 suggests strong correlation among items with their respective factor denoting high internal consistency (34).

Test–retest reliability

A subsample of young adults (n = 31), who had previously agreed to be contacted for future research, completed the WREQ for a second time within an average of 4.1 weeks (range 2.9–6.4 weeks) between test dates. Test–retest reliability was assessed at the subscale level via intraclass correlations (ICC) and paired sample t-tests. ICC values above 0.75 denote excellent reliability, values between 0.40 and 0.75 represents fair-to-good reliability, and values below 0.40 are suggestive of poor reliability (35). Analyses were conducted using SPSS version 16.0 (SPSS, Chicago, IL).

Multi-group factorial invariance

The factorial invariance of the WREQ between population subgroups was examined by fitting and comparing a sequence of CFA models using Mplus Version 5.21 (28). Tests of population factorial invariance determine the existence of test bias, pertaining to the consistency of measurement properties across subgroups. More specifically, factorial invariance supports that the associations between the indicators (questionnaire items) and the factors (subscales) are equivalent across groups. Analyses began with an unconstrained (i.e., configural) model and progressed toward more restricted (nested) models to evaluate the tenability of each sequentially placed constraint. The configural model was initially fit in each subgroup, functioning as a prerequisite or baseline model for additional invariance tests. Next, sequential model constraints within nested models were imposed, examining invariance of model form, factor loadings, and indicator intercepts across gender, age, race, and BMI subgroups. Invariant model form indicates equality in the number of latent factors and pattern of indicator-factor loadings. Tests of invariant factor loadings and indicator intercepts determine whether the latent factors have the same meaning and structure for different groups of participants.

The tenability of invariance at each level of model constraint was determined using the change in (Δ) degrees of freedom (df) of the parent model to the constrained model (30) and minimal ΔCFI ($CFI_{\text{constrained model}} - CFI_{\text{unconstrained model}}$ ≤0.01) (36). Like the χ^2 statistic,

$\Delta\chi^2$ is also susceptible to sample size and may reject null hypotheses due to trivial model differences. Based on results of a Monte Carlo simulation study, Cheung and Rensvold (36) determined that ΔCFI ≤0.01 is superior to $\Delta\chi^2$ and among the best change index for determining the consequences of nested model constraints. Although further validation of this report is warranted, recommendations provide the most justifiable evidence to date; hence, the current study emphasized ΔCFI as the final determinant of factorial invariance.

Criterion-related validity

Weight-related variables including current BMI, weight control status, and weight change rate were used to validate the subscales of the WREQ. Analyses were conducted using hierarchical multivariate regression to explore associations between the WREQ subscales and current BMI as well as weight change rate to control for significant correlations among the subscales. For all the models, analyses were repeated to control for additional confounders including gender, age, and race. Multivariate regression analyses for weight change rate were conducted separately for young adults and adults based on the difference in the length time on which weight change rate was calculated. Multivariate analysis of covariance was used to explore differences in WREQ subscale scores by weight control status. Again, gender, age, and race were included in the model as covariates. Group differences were evaluated using pairwise comparisons with Bonferroni adjustments. All construct validation analyses were conducted using SPSS version 16.0 (SPSS). Significance was set at $P < 0.05$.

Results

The final study sample consisted of 257 young adults (29.7% males) and 334 adults (45.5% males). The adults had a mean age of 46.0 ± 12.9 years and a BMI 27.6 ± 6.3 kg/m^2. The young adults had a mean age of 19.5 ± 1.6 years and a BMI 23.3 ± 5.1 kg/m^2. Based on the World Health Organization standards, 7.2% of young adults and 1.8% of adults were underweight (BMI <18.5 kg/m^2), 67.5% of young adults and 33.4% of adults were of normal weight (BMI = 18.5–24.9 kg/m^2), 17.4% of young adults and 36.7% of adults were overweight (BMI = 25–29.9 kg/m^2), and 7.9% of young adults and 28.0% of adults were categorized as obese (BMI >29.9 kg/m^2). A majority of the participants were of Asian/Asian mix (42.2%), 23.2% were white, and 17.9% were at least part-Native Hawaiian/Pacific Islander. The remaining participants (16.4%) were of other mixed race. Normative scores and Cronbach's α coefficients for the WREQ subscales for the whole sample and by group are presented as Table 1. Group differences in some WREQ subscale scores were observed by gender, age, and weight status, but not by race.

Factorial validity and internal consistency

Figure 1 illustrates the hypothesized correlated four-factor structure for the WREQ. The first item of each latent factor was used as an indicator for its corresponding latent factor by constraining the parameter to a value of one. The tested model contained no double loading indicators and all measurement error was presumed to be uncorrelated. For the total sample, all freely estimated unstandardized parameters were statistically significant ($P \leq 0.001$). Inspection of residual variances and modification indices indicated no ill fits within the solution and standardized parameter estimates indicated that the items were strongly related to their corresponding latent factors (R^2 = 0.216–0.820). The four latent constructs were moderately correlated (r = 0.10–0.48), and factor determinacy coefficients met recommended standards (34). Model fit statistics and factor determinacy coefficients for the total sample and each subgroup are presented in Table 2.

Test–retest reliability

All the WREQ subscales had reliability coefficients above standard criteria suggesting good test–retest reliability. The ICC for the WREQ met the criteria for excellent reliability compensatory restraint (ICC = 0.890); routine restraint (ICC = 0.864); external eating (ICC = 0.849); and emotional eating (ICC = 0.932). The test–retest coefficients via paired *t*-test were also acceptable: compensatory restraint ($r = 0.804$); routine restraint ($r = 0.760$); external eating ($r = 0.742$); and emotional eating ($r = 0.857$).

Population factorial invariance

Tests of equal model form across groups was acceptable between males and females (χ^2 = 400.58, df = 208, CFI = 0.94, SRMR = 0.05), Asians/Asian mix and whites (χ^2 = 339.65, df = 208, CFI = 0.93, SRMR = 0.06), young adults and adults (χ^2 = 356.63, df = 208, CFI = 0.96, SRMR = 0.05), and underweight/normal weight and overweight/obese subgroups (χ^2 = 379.30, df = 208, CFI = 0.96, SRMR = 0.05). The two nested comparisons indicated equivalent actor loadings and indicator intercepts were invariant between genders ($\Delta\chi^2$ = 16.71, Δdf = 12, $P > 0.05$; ΔCFI = 0.00), races ($\Delta\chi^2$ = 14.30, Δdf = 12, $P > 0.05$; ΔCFI = 0.00), age groups ($\Delta\chi^2$ = 7.22, Δdf = 12, $P > 0.05$; ΔCFI = 0.00), and BMI groups ($\Delta\chi^2$ = 12.25, Δdf = 12, $P > 0.05$; ΔCFI = 0.00).

Criterion-related validity

Higher emotional eating scores were positively associated with current BMI for the whole sample by multivariate regression analyses that included compensatory restraint, routine restraint, and external eating. This association remained significant after accounting for gender, age, and race (Table 3).

By weight control status, those who identified themselves as currently dieting had highest routine restraint scores (Table 4). Compensatory restraint and emotional eating were similarly high in dieters and those who were trying to avoid weight gain. Differences remained significant after accounting for gender, age, race, and current BMI.

With regards to weight change rate, a total of 22% of young adults reported losing more than 1 kg over a 5-month period and 35% reported gaining more than 1 kg (range 1.5–11.4 kg). After a period of 25.0 ± 12.9 years (range 4–60 years), ~10% of adults reported weighing less than they did at 21 years of age and 85% reported being at least 1 kg (range 1.5–76.8 kg) more than they were at 21 years of age. Results of the multivariate regression analyses (Table 5) demonstrated that higher emotional eating scores were associated with greater weight change rates (weight gain) among young adults as well as in adults after accounting for intercorrelations among the subscales. For adults, compensatory restraint was also found to be negatively associated weight change rate. These findings remained significant after accounting for gender, age, and race.

Discussion

Construct validation conducted in this study provide strong support for the four-factor structure of the WREQ, which is represented by two subscales of dietary restraint (compensatory and RRroutine restraint), external eating, and emotional eating with strong internal consistency and test–retest coefficients. Factorial invariance testing provided no evidence that the factor structure, factor loadings, and indictor intercepts were different between males and females, young adults and adults, whites and Asians, or those categorized as under/normal weight and overweight/obese. As hypothesized, criterion-related validation analyses confirmed that the WREQ subscales were associated with weight-related variables including current BMI, weight change rates, and/or weight control status, exception for external eating. Collectively, these analyses demonstrate the WREQ is acceptable for use in research studies exploring the cognitive and psychological aspects of eating behavior as they relate to body weight in diverse populations.

The psychometric evaluation of the WREQ confirmed a four-factor structure of the WREQ, which parallels both fundamental and theoretical expectations (6–8) and subsequent research that suggests that dietary restraint is not a homogeneous scale (17). The confirmation of factorial invariance of the WREQ across multiple subgroups gives researchers the confidence that their research findings are interpretable, particularly in populations that incorporate individuals of various age, gender, or race/ethnicity (Asians and whites). Previous research attempting to replicate the factor structure and reliability statistics for similar questionnaires has reported mixed findings among adult populations categorized by gender (11,13,37), and/or weight status (10,11,13), as well as by clinical and nonclinical status (eating disordered or diabetic) (10,13). Rarely, if at all, have studies been conducted to explore the consistency of the psychometric properties of questionnaires by age or race groups and only two studies were found that utilized factorial invariance methods to evaluate similar psychometric instruments (37,38). Prior to utilizing a psychometric instrument in diverse populations, it should be evaluated for use in subgroups of the population, particularly if those subgroups are different from which the questionnaire was developed. Beyond test–retest reliability or internal consistency statistics (e.g., Cronbach's α), most often used to support the use of a questionnaire in a specific population, it is important to confirm that the constructs are similarly measured across populations. Techniques, such as measurement invariance, are optimal methods for validating questionnaires across subgroups of populations. Confirmation of factorial invariance ensures that scales/subscales have the same general meaning between groups. A lack of factorial invariance could ultimately lead to biased mean values and/or systematic errors in the interpretation of research findings.

As hypothesized, criterion-related validation analyses conducted in this study provides support for the characterization of four subscales as weight-related aspects of eating behavior. Specifically, compensatory restraint was found to be protective against weight gain from 21 years of age throughout adulthood suggesting that individuals who purposefully eat less before or after a known period of overconsumption are less prone to excessive weight gains compared to people who do not self-regulate their energy intake. Routine restraint scores, however, were not associated with long-term changes in weight despite being highest among dieters. These findings are consistent with observations that the flexible component of dietary restraint attenuates 20-year weight gains in women between the ages of 20 and 50 years (39) whereas, more rigid approaches to weight control may not promote weight control over moderate to long time periods (39,40). It is unknown from this study, if compensatory restraint is a successful approach for dieting and would require longitudinal examinations. However, it is likely that adopting a compensatory (or flexible) approach to eating as a habitual "lifestyle" change, rather than a short-term weight loss practice, will prove to be most successful for long-term weight control. Unlike external eating, the emotional eating subscale showed robust and independent positive associations with weight and weight change over various observation periods in young adults and adults. Similarly, with other eating behavior questionnaires, this association appears to be robust across various populations and various ages (23,39,41) suggesting that the WREQ emotional eating subscale varies very little from previously validated measures.

In this study, external eating was not found to be associated with current weight, weight control status, or weight change rates. While this was not entirely unexpected, it is generally accepted that an increased susceptibility to external cues to eat could lead to excessive caloric intake (3,42). Yet, few researchers have reported independent associations between external eating and weight or weight change using explicit measures (43). Rather, implicit tests (e.g., Stroop testing) consistently show that overweight and obese individual have greater reactivity to food cues (measured by eye movement and brain activation) and have higher external eating scores (44,45). There are two probable reasons for a lack of association between external and weight change. The first may be explained by the observation that most individuals are unaware of the external influences that motivate eating and therefore, are less likely to be able to accurately respond to questions like those in the WREQ or similar questionnaires (3). The second may be related to the circumstances or amounts of foods eaten, when externally cued to eat. For instance, one may be externally cued to eat, but the resulting eating event may occur simultaneously with an internally motivated cue to eat (physiological hunger) or may adequately delay a subsequent eating episode resulting in a relatively balanced energy intake for the day (46). This does not necessarily suggest that external eating is not associated with weight-related variables. It is hypothesized that the joint influence of external eating and a lack of sensitivity to internal cues may elucidate the role of external eating as a weight-related aspect of eating.

There are multiple strengths and limitations of this research. This was the first study of its kind to evaluate the properties of an eating behavior questionnaire in large and diverse population. The use of structural equation modeling to perform CFA and tests of factorial invariance greatly enhances the confidence in the findings given that this statistical method takes into account measurement error and provides fit indices for tested models, which can be objectively interpreted. Also, with regards to construct validation, weight control status and weight change over 5 months (moderate term) and an average of 25 years (long term) in addition to weight status (current BMI) were utilized. These variables provide a more comprehensive measure of short- (e.g., currently dieting) to long-term weight control (25-year weight change) as they relate to eating behaviors, than does BMI alone, allowing for a broader understanding of dietary restraint, external eating, and emotional eating. For this, we relied on retrospective, self-report of weight that could be recalled incorrectly, particularly among older adults recalling their weight after more than 25 years. However, there is evidence that weight earlier in life can be recalled with accuracy in older subjects (47), the Nurses Health Study (48), a follow-up of American adults (49), and the Newton Girls' Study (50). Other limitations are based on the sampled population. Specifically, there were very few Latinos in the sampled population, suggesting that additional invariance testing would be recommended in this subgroup. However, based on the robust findings in this study, it is likely that factorial invariance would be confirmed in this population if tested. Additionally, these findings cannot be generalized to clinical populations, particular eating disordered populations. Again, further psychometric and invariance testing would be recommended.

To summarize, this is the first study to extensively test the psychometric properties, measurement invariance, and construct validation of an eating behavior questionnaire in such a diverse population. Results demonstrate that the WREQ measures four, internally-reliable and meaningful constructs of eating behavior with factorial invariance between males and females, young adults and adults, under/normal weight and overweight/obese, and whites and Asians/Asian-mix. Criterion-related validation analyses indicated that the WREQ subscale are associated with weight-related variables in a manner similar to other eating behavior questionnaires and may have predictive ability, particularly over time periods of modest length (e.g., several months). Therefore, the WREQ can be used as a reliable and valid measure of cognitive and psychological eating behaviors in a variety of research populations.

Acknowledgments

The authors are supported by postdoctoral fellowships on National Cancer Institute grant R25CA 90956.

REFERENCES

1. Ogden CL, Carroll MD, Curtin LR, Lamb MM, Flegal KM. Prevalence of high body mass index in US children and adolescents, 2007–2008. JAMA. 2010; 303:242–249. [PubMed: 20071470]
2. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of overweight and obesity in the United States, 1999–2004. JAMA. 2006; 295:1549–1555. [PubMed: 16595758]
3. Wansink, B. Mindless Eating: Why We Eat More Than We Think. New York: Bantam Books; 2006.
4. Berthoud HR. Interactions between the “cognitive” and “metabolic” brain in the control of food intake. Physiol Behav. 2007; 91:486–498. [PubMed: 17307205]
5. Rutters F, Nieuwenhuizen AG, Lemmens SG, Born JM, Westerterp-Plantenga MS. Acute stress-related changes in eating in the absence of hunger. Obesity (Silver Spring). 2009; 17:72–77. [PubMed: 18997672]
6. Kaplan HI, Kaplan HS. The psychosomatic concept of obesity. J Nerv Ment Dis. 1957; 125:181–201. [PubMed: 13481715]
7. Schachter S, Gross LP. Manipulated time and eating behavior. J Pers Soc Psychol. 1968; 10:98–106. [PubMed: 5725908]
8. Herman CP, Mack D. Restrained and unrestrained eating. J Pers. 1975; 43:647–660. [PubMed: 1206453]
9. Herman, CP.; Polivy, J. Restrained eating. In: Stunkard, AJ., editor. Obesity. Philadelphia: W. B. Saunders; 1980. p. 208-225.
10. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. J Psychosom Res. 1985; 29:71–83. [PubMed: 3981480]
11. van Strien T, Frijters JER, Bergers GPA, Defares PB. The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. Int J Eat Disord. 1986; 5:295–315.
12. de Lauzon B, Romon M, Deschamps V, et al. Fleurbaix Laventie Ville Sante Study Group. The Three-Factor Eating Questionnaire-R18 is able to distinguish among different eating patterns in a general population. J Nutr. 2004; 134:2372–2380. [PubMed: 15333731]
13. Cappelleri JC, Bushmakin AG, Gerber RA, et al. Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants. Int J Obes (Lond). 2009; 33:611–620. [PubMed: 19399021]
14. Karlsson J, Persson LO, Sjöström L, Sullivan M. Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. Int J Obes Relat Metab Disord. 2000; 24:1715–1725. [PubMed: 11126230]
15. Lowe MR, Butryn ML, Didie ER, et al. The Power of Food Scale. A new measure of the psychological influence of the food environment. Appetite. 2009; 53:114–118. [PubMed: 19500623]
16. Cappelleri JC, Bushmakin AG, Gerber RA, et al. Evaluating the Power of Food Scale in obese subjects and a general sample of individuals: development and measurement properties. Int J Obes (Lond). 2009; 33:913–922. [PubMed: 19506564]
17. Westenhoefer J. Dietary restraint and disinhibition: is restraint a homogeneous construct? Appetite. 1991; 16:45–55. [PubMed: 2018403]
18. Bond MJ, McDowell AJ, Wilkinson JY. The measurement of dietary restraint, disinhibition and hunger: an examination of the factor structure of the Three Factor Eating Questionnaire (TFEQ). Int J Obes Relat Metab Disord. 2001; 25:900–906. [PubMed: 11439306]
19. Stice E, Fisher M, Lowe MR. Are dietary restraint scales valid measures of acute dietary restriction? Unobtrusive observational data suggest not. Psychol Assess. 2004; 16:51–59. [PubMed: 15023092]
20. Stice E, Cooper JA, Schoeller DA, Tappe K, Lowe MR. Are dietary restraint scales valid measures of moderate- to long-term dietary restriction? Objective biological and behavioral data suggest not. Psychol Assess. 2007; 19:449–458. [PubMed: 18085937]

21. Stice E, Sysko R, Roberto CA, Allison S. Are dietary restraint scales valid measures of dietary restriction? Additional objective behavioral and biological data suggest not. Appetite. 2010; 54:331–339. [PubMed: 20006662]

22. Keskitalo K, Tuorila H, Spector TD, et al. The Three-Factor Eating Questionnaire, body mass index, and responses to sweet and salty fatty foods: a twin study of genetic and environmental associations. Am J Clin Nutr. 2008; 88:263–271. [PubMed: 18689360]

23. Sung J, Lee K, Song YM. Relationship of eating behavior to long-term weight change and body mass index: the Healthy Twin study. Eat Weight Disord. 2009; 14:e98–e105. [PubMed: 19934643]

24. Schembre S, Greene G, Melanson K. Development and validation of a weight-related eating questionnaire. Eat Behav. 2009; 10:119–124. [PubMed: 19447354]

25. Westenhoefer J, Stunkard AJ, Pudel V. Validation of the flexible and rigid control dimensions of dietary restraint. Int J Eat Disord. 1999; 26:53–64. [PubMed: 10349584]

26. Timko CA, Perone J. Rigid and flexible control of eating behavior in a college population. Eat Behav. 2005; 6:119–125. [PubMed: 15598598]

27. Kolonel LN, Henderson BE, Hankin JH, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol. 2000; 151:346–357. [PubMed: 10695593]

28. Muthen, LK.; Muthen, BO. Mplus User's Guide. 5th edn. Los Angeles, CA: Muthen & Muthen; 2007.

29. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990; 107:238–246. [PubMed: 2320703]

30. Bollen, KA. Structural Equations with Latent Variables. New York: Wiley; 1989.

31. Joreskog, KG. Testing structural equation models. In: Bollen, KA., editor. Testing Structural Equation Models. Newbury Park: Sage Publications; 1993. p. 294-316.

32. Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. Testing Structural Equation Models. Newbury Park, CA: Sage; 1993. p. 136-162.

33. Hu L-T, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999; 6:1–55.

34. Gorsuch, RL. Factor Analysis. 2nd edn.. Hillsdale, NJ: L. Erlbaum Associates; 1983.

35. Fleiss, JL. The Design and Analysis of Clinical Experiments. New York: John Wiley and Sons; 1986.

36. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling. 2002; 9:233–255.

37. Boerner LM, Spillane NS, Anderson KG, Smith GT. Similarities and differences between women and men on eating disorder risk factors and symptom measures. Eat Behav. 2004; 5:209–222. [PubMed: 15135333]

38. van Strien T, Oosterveld P. The children's DEBQ for assessment of restrained, emotional, and external eating in 7- to 12-year-old children. Int J Eat Disord. 2008; 41:72–81. [PubMed: 17634965]

39. Hays NP, Roberts SB. Aspects of eating behaviors "disinhibition" and "restraint" are related to weight gain and BMI in women. Obesity (Silver Spring). 2008; 16:52–58. [PubMed: 18223612]

40. Tiggemann M. Dietary restraint and self-esteem as predictors of weight gain over an 8-year time period. Eat Behav. 2004; 5:251–259. [PubMed: 15135337]

41. Braet C, Claus L, Goossens L, et al. Differences in eating style between overweight and normal-weight youngsters. J Health Psychol. 2008; 13:733–743. [PubMed: 18697886]

42. Herman CP, Polivy J. External cues in the control of food intake in humans: the sensory-normative distinction. Physiol Behav. 2008; 94:722–728. [PubMed: 18499202]

43. Anschutz DJ, Van Strien T, Van De Ven MO, Engels RC. Eating styles and energy intake in young women. Appetite. 2009; 53:119–122. [PubMed: 19481836]

44. Nijs IM, Franken IH, Muris P. Enhanced processing of food-related pictures in female external eaters. Appetite. 2009; 53:376–383. [PubMed: 19647769]

45. Nijs IM, Muris P, Euser AS, Franken IH. Differences in attention to food and food intake between overweight/obese and normal-weight females under conditions of hunger and satiety. Appetite. 2010; 54:243–254. [PubMed: 19922752]

46. McKiernan F, Hollis JH, Mattes RD. Short-term dietary compensation in free-living adults. Physiol Behav. 2008; 93:975–983. [PubMed: 18261752]

47. Must A, Willett WC, Dietz WH. Remote recall of childhood height, weight, and body build by elderly subjects. Am J Epidemiol. 1993; 138:56–64. [PubMed: 8333427]

48. Troy LM, Hunter DJ, Manson JE, et al. The validity of recalled weight among younger women. Int J Obes Relat Metab Disord. 1995; 19:570–572. [PubMed: 7489028]

49. Perry GS, Byers TE, Mokdad AH, Serdula MK, Williamson DF. The validity of self-reports of past body weights by U.S. adults. Epidemiology. 1995; 6:61–66. [PubMed: 7888448]

50. Must A, Phillips SM, Naumova EN, et al. Recall of early menstrual history and menarcheal body size: after 30 years, how well do women remember? Am J Epidemiol. 2002; 155:672–679. [PubMed: 11914195]

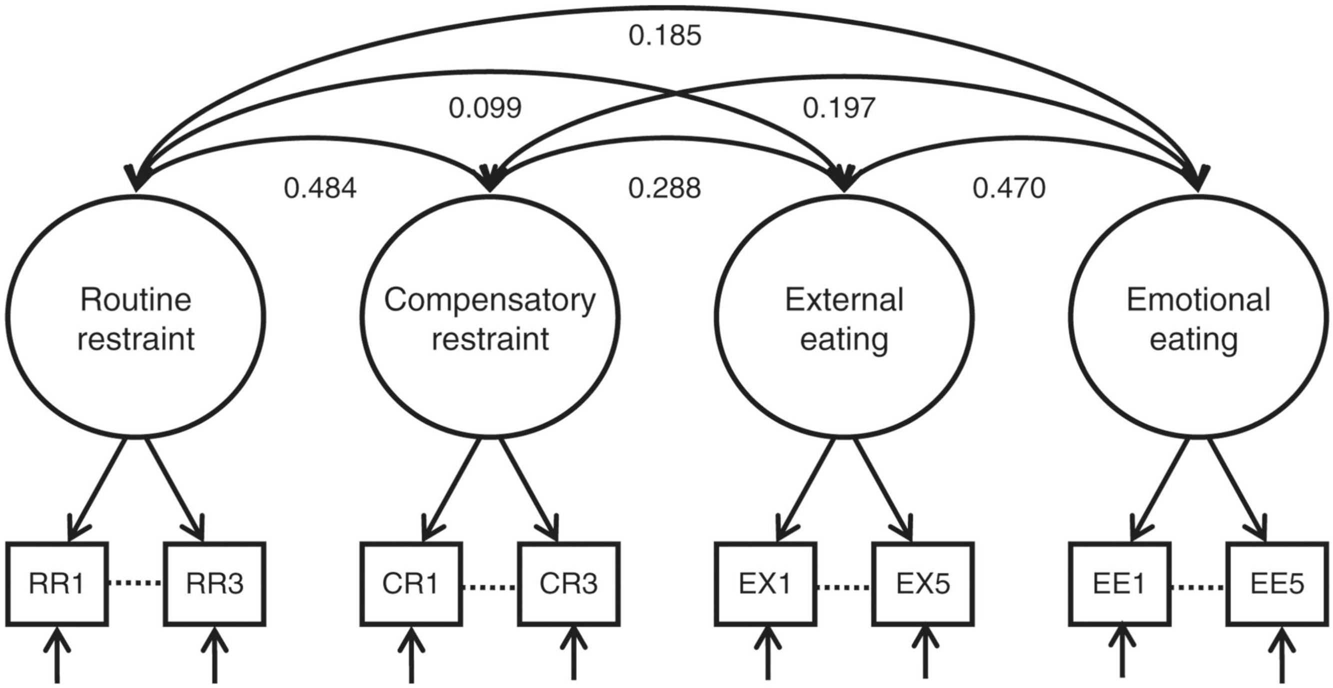

Figure 1.
Four, correlated factors model of the Weight-Related Eating Questionnaire with Pearson correlation coefficients. RR1–RR3 represent item indicators for routine restraint, CR1–CR3 represent indicator items for compensatory restraint, EX1–EX5 represent indicator items for susceptibility to external cues, and EE1–EE5 represent indicator items for emotional eating.

Table 1

Descriptive statistics for the WREQ subscales by select subgroups

	All subjects (n = 621)				
	Mean	s.d.	s.e.	Range	α
Dietary restraint	2.22	0.72	0.03	1–4.83	0.777
Compensatory restraint	2.52	0.92	0.04	1–5	0.667
Routine restraint	1.92	0.75	0.03	1–4.67	0.749
External eating	2.69	0.86	0.03	1–5	0.800
Emotional eating	2.00	1.00	0.04	1–5	0.908

	Males (n = 242)					Females (n = 379)						
	Mean	s.d.	s.e.	Range	α	Mean	s.d.	s.e.	Range	α	$P_{difference}$	ES
Dietary restraint	2.14	0.70	0.05	1–4.5	0.768	2.27	0.73	0.04	1–4.83	0.782	0.030	0.005
Compensatory restraint	2.42	0.89	0.06	1–5	0.639	2.59	0.94	0.05	1–5	0.683	0.019	0.005
Routine restraint	1.87	0.74	0.05	1–4.33	0.748	1.95	0.75	0.04	1–4.67	0.750	0.196	0.003
External eating	2.65	0.87	0.06	1–5	0.802	2.72	0.86	0.04	1–5	0.800	0.338	0.013
Emotional eating	1.80	0.94	0.06	1–5	0.913	2.13	1.02	0.05	1–5	0.902	<0.001	0.006

	Young adults (n = 266)					Adults (n = 334)						
	Mean	s.d.	s.e.	Range	α	Mean	s.d.	s.e.	Range	α	$P_{difference}$	ES
Dietary restraint	2.16	0.75	0.05	1–4.83	0.801	2.27	0.70	0.04	1–4.67	0.762	0.075	0.008
Compensatory restraint	2.44	0.95	0.06	1–5	0.685	2.58	0.89	0.05	1–5	0.670	0.075	0.009
Routine restraint	1.88	0.76	0.05	1–4.67	0.768	1.96	0.75	0.04	1–4.67	0.728	0.213	0.003
External eating	2.80	0.83	0.05	1–5	0.790	2.61	0.89	0.05	1–5	0.813	0.005	0.001
Emotional eating	1.91	0.93	0.06	1–4.6	0.895	2.07	1.06	0.06	1–5	0.916	0.059	0.024

	White (n = 144)					Asian/Asian mix (n = 264)						
	Mean	s.d.	s.e.	Range	α	Mean	s.d.	s.e.	Range	α	$P_{difference}$	ES
Dietary restraint	2.31	0.76	0.06	1–4.5	0.835	2.21	0.67	0.04	1–4	0.737	0.202	0.004

	All subjects (n = 621)											
	Mean	**s.d.**	**s.e.**	**Range**	**α**							
Compensatory restraint	2.60	0.92	0.08	1–4.67	0.697	2.54	0.89	0.05	1–4.67	0.653	0.537	0.001
Routine restraint	2.02	0.77	0.06	1–4.33	0.804	1.90	0.71	0.04	1–4	0.726	0.091	0.007
External eating	2.65	0.86	0.07	1–4.6	0.802	2.76	0.82	0.05	1–5	0.784	0.233	0.003
Emotional eating	2.18	1.10	0.09	1–5	0.917	1.98	0.96	0.06	1–4.6	0.902	0.056	0.010
	Under/normal weight (n = 326)					**Overweight/obese (n = 292)**						
	Mean	**s.d.**	**s.e.**	**Range**	**α**	**Mean**	**s.d.**	**s.e.**	**Range**	**α**	$P_{difference}$	**ES**
Dietary restraint	2.19	0.76	0.04	1–4.83	0.808	2.26	0.67	0.04	1–4.5	0.734	0.188	0.003
Compensatory restraint	2.53	0.97	0.05	1–5	0.729	2.52	0.86	0.05	1–4.67	0.580	0.947	0.000
Routine restraint	1.85	0.78	0.04	1–4.67	0.785	2.00	0.70	0.04	1–4.33	0.695	0.009	0.011
External eating	2.67	0.88	0.05	1–5	0.816	2.72	0.85	0.05	1–5	0.784	0.469	0.001
Emotional eating	1.80	0.91	0.05	1–4.6	0.896	2.22	1.07	0.06	1–5	0.914	<0.001	0.043

α, Cronbach's **α**; ES, effect size (partial eta squared); WREQ, Weight-Related Eating Questionnaire.

Table 2

Fit indices and reliabilities by gender, age, race, and weight status groups

Subgroup	*n*	χ^2 (df)	CFI	SRMR	RMSEA	CFit	TLI	Factor score determinacy			
								Compensatory restraint	Routine restraint	External eating	Emotional eating
Total	621	233.41 (98)	0.97	0.03	0.05	0.72	0.96	0.88	0.87	0.91	0.96
Males	242	199.13 (98)	0.94	0.05	0.07	0.03	0.92	0.88	0.87	0.91	0.96
Females	242	170.58 (98)	0.95	0.04	0.06	0.26	0.94	0.90	0.88	0.91	0.95
Young adults	266	155.80 (98)	0.97	0.04	0.05	0.62	0.96	0.91	0.90	0.91	0.95
Adults	266	177.33 (98)	0.96	0.04	0.06	0.25	0.95	0.89	0.88	0.93	0.97
White	144	162.41 (98)	0.94	0.05	0.07	0.06	0.93	0.93	0.90	0.91	0.97
Asian/Asian-mix	144	139.75 (98)	0.95	0.06	0.05	0.35	0.94	0.87	0.86	0.92	0.95
Underweight/normal weight	292	187.54 (98)	0.96	0.04	0.06	0.20	0.95	0.90	0.92	0.92	0.96
Overweight/obese	292	154.59 (98)	0.97	0.05	0.04	0.75	0.96	0.88	0.85	0.91	0.96

CFI, comparative fit index; CFit, close fit (probability RMSEA ≤0.05); df, degrees of freedom; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker–Lewis index.

Table 3

Multivariate regression with current BMI (n = 597)

	r^a	*B*	s.e.	β	*t*-value	*P* value	Adj. *P* value[b]
Constant		23.418	0.997		23.48	<0.001	0.007
Compensatory restraint	0.016	−0.390	0.319	−0.058	−1.22	0.222	0.215
Routine restraint	0.093	0.631	0.376	0.077	1.68	0.094	0.204
External eating	0.048	−0.478	0.328	−0.067	−1.46	0.145	0.550
Emotional eating	0.244	1.674	0.278	0.272	6.01	<0.001	<0.001

B, unstandardized coefficients, β, standardized coefficients.

[a]Pearson correlation coefficients.

[b]Adjusted model includes gender ($P < 0.001$), age ($P < 0.001$), race ($P = 0.210$).

Table 4

Mean WREQ scores by weight control status responses

	I am trying to gain weight ($n = 32$)	I am not concerned with my weight ($n = 139$)	I am trying to avoid weight gain ($n = 354$)	I am dieting to lose weight ($n = 70$)	Kruskal–Wallis *P* value	Adjusted *P* value[a]
BMI	21.26 ± 2.81^{A}	24.33 ± 5.65^{B}	25.90 ± 6.29^{B}	29.36 ± 5.78^{C}	<0.001	<0.001
Compensatory restraint	1.69 ± 0.75^{A}	2.21 ± 0.83^{A}	2.64 ± 0.88^{B}	2.90 ± 0.99^{B}	<0.001	<0.001
Routine restraint	1.31 ± 0.51^{A}	1.41 ± 0.51^{A}	2.06 ± 0.68^{B}	2.51 ± 0.84^{C}	<0.001	<0.001
External eating	2.52 ± 0.84	2.69 ± 0.88	2.70 ± 0.87	2.76 ± 0.87	0.508	0.289
Emotional eating	1.39 ± 0.49^{A}	1.88 ± 0.97^{A}	$2.00 \pm 0.95^{A,B}$	2.48 ± 1.27^{B}	<0.001	<0.001

Follow-up analyses for differences between categories were conducted using multivariate analysis of covariance; values with different superscript letters are significantly different at $P < 0.05$ with Bonferroni adjustments for multiple comparisons.

WREQ, Weight-Related Eating Questionnaire.

[a] Adjusted multivariate model included gender ($P < 0.001$), age ($P < 0.001$), race ($P = 0.243$)

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Multivariate regression with weight change rate by age categories

Young adults (n = 261)	r[a]	*B*	s.e.	β	*t*-value	*P* value	Adj. *P* value[b]
Constant		−1.177	1.965		−0.60	0.550	0.007
Compensatory restraint	−0.005	−0.071	0.625	−0.009	−0.11	0.910	0.999
Routine restraint	−0.062	−1.174	0.775	−0.115	−1.52	0.131	0.156
External eating	0.134	0.696	0.692	0.074	1.01	0.315	0.507
Emotional eating	0.151	1.267	0.619	0.152	2.05	0.042	0.017
Adults (n = 312)	**r[a]**	***B***	**s.e.**	**β**	***t*-value**	***P* value**	**Adj. *P* value[c]**
Constant		0.496	0.158		3.15	0.002	<0.001
Compensatory restraint	−0.106	−0.105	0.051	−0.133	−2.07	0.039	0.025
Routine restraint	−0.061	−0.020	0.570	−0.022	−0.36	0.723	0.624
External eating	0.095	0.028	0.051	0.035	0.54	0.587	0.958
Emotional eating	0.215	0.144	0.041	0.217	3.50	0.001	0.001

B, unstandardized coefficients, β, standardized coefficients.

[a]Pearson correlation coefficients.

[b]Adjusted model for young adults includes gender (P = 0.040), age (P = 0.008), race (P = 0.803).

[c]Adjusted model for adults includes gender (P = 0.789), age (P < 0.001), race (P = 0.764).