Published in final edited form as:

Body Image. 2022 June ; 41: 97–108. doi:10.1016/j.bodyim.2022.01.015.

Measurement Invariance of Body Image Measures by Age, Gender, Sexual Orientation, Race, Weight Status, and Age: The U.S. Body Project I

Vivienne M. Hazzard[a], **Lauren M. Schaefer**[a], **J. Kevin Thompson**[b], **Stuart B. Murray**[c], **David A. Frederick**[d]

[a.]Sanford Center for Bio-behavioral Research, Fargo, ND, USA

[b.]Department of Psychology, University of South Florida, FL, USA

[c.]Department of Psychiatry & Behavioral Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

[d.]Crean College of Health and Behavioral Sciences, Chapman University, Orange, CA, USA

Abstract

Despite growing interest in comparing body image experiences across diverse groups, limited work has examined whether body image measures operate similarly across different populations, raising important questions about the appropriateness of comparing scale means across demographic groups. This study employed measurement invariance testing to evaluate whether such comparisons are appropriate with existing body image measures. Specifically, multi-group confirmatory factor analysis was conducted using a community sample of 11,620 men and women to test increasing levels of invariance (configural, metric, scalar) across five key demographic variables (age group, gender, sexual orientation, race, weight status) for five commonly used body image measures (the Sociocultural Attitudes Towards Appearance Questionnaire-4, the Body Surveillance subscale of the Objectified Body Consciousness Scale, the Appearance Evaluation and Overweight Preoccupation subscales of the Multidimensional Body-Self Relations Questionnaire, and the Body Image Quality of Life Inventory). Results provided evidence of scalar (i.e., strong) invariance for all five measures across age, gender, sexual orientation, race, and weight status groups, indicating that the latent factors captured by these measures have the same meaning across demographic groups. Findings therefore support the comparison of scale/subscale means across multiple demographic groups for these body image measures.

Corresponding author: Dr. Vivienne M. Hazzard at viviennehazzard@gmail.com or Dr. David A. Frederick at enderflies1@aol.com.

Author Statement

The last author completed the conceptualization and methodology. The first author took the lead on writing and data analyses. All authors engaged in the writing, and made suggestions for interpreting and enhancing the formal analysis. The last authors two authors engaged in supervision. The last author was involved with funding acquisition (grant PI).

Declaration of interest statement: The authors have no conflicts of interest or financial conflicts relevant to this article to disclose.

Conflict of Interest

There are no conflicts of interest.

Keywords

Body Image; Psychometrics; Gender; Race; Sexual Orientation; Body Mass Index

1. Introduction

Body image is a multidimensional construct comprised of the thoughts, emotions, and behaviors related to an individual's appearance (Cash & Smolak, 2011). Given the complex nature of body image, numerous measures have been developed to assess the individual aspects of body image that are differentially related to proposed outcomes. For example, researchers have created measures to assess satisfaction or dissatisfaction with appearance (Brown et al., 1990), concern regarding one's weight (Brown et al., 1990), the degree to which one's body image impacts their daily lives and emotional wellbeing (Cash & Fleming, 2002), degree of body self-monitoring (McKinley & Hyde, 1996), perceived social pressures regarding appearance (Schaefer et al., 2015), and internalization of socially-prescribed appearance ideals (Schaefer et al., 2015).

Within the field of body image, there has been a long-standing interest in examining body image experiences and their correlates in different populations to identify convergent and divergent processes across people of different ages, genders, sexual orientations, races, and weight statuses (Fallon et al., 2014; Frederick & Essayli, 2016; Frederick, Kelly, et al., 2016; Frederick, Sandhu, et al., 2016; Morrison et al., 2004; Peplau et al., 2009; Roberts et al., 2006). This literature identifies potentially important differences in body image experiences across groups. For example, studies have found higher scores on measures of appearance evaluation and overweight preoccupation among women compared to men (Frederick & Forbes et al., 2007), lower levels of body surveillance among Black women compared to White women (Schaefer et al., 2018), and greater body dissatisfaction among people with higher body masses (Forbes & Frederick, 2008; Frederick et al., 2020). Age, on the other hand, has generally been unrelated or only weakly related to body satisfaction (Tiggemann, 2004). Regarding sexual orientation, research has consistently found that gay men report greater body dissatisfaction (Frederick & Essayli, 2016) and that sexual minority men report more disordered eating (Murray et al., 2017), whereas sexual orientation differences among women are generally small or negligible (He et al., 2020).

We examined measurement invariance across groups in five key demographic categories that are known to be associated with life outcomes, people's experiences, and body image outcomes: age, gender, sexual orientation, race, and weight status. Conclusions drawn from past research are dependent on there being invariance in responses to items in scales across these different groups. It could be the case, however, that there are group differences in an underlying construct, despite groups have identical means on scales. For example, appearance evaluation assesses how attractive people feel they are and satisfaction with appearance (Brown et al., 1990), and body surveillance measures how much people monitor how they appear to others and how concerned they are with how they versus what their bodies can physically do (McKinley & Hyde, 1996). Both of these measures contain items about clothing (Body Surveillance: "I often worry about whether the clothes I am wearing make me look good;" "I think it is more important that my clothes are comfortable than whether they look good on me;" Appearance Evaluation: "I like the way my clothes fit me"). Women's fashion often emphasizes physical attributes and can be tight fitting, so these items might apply more so to women. Women who score high on the body surveillance measure might attain these high scores by expressing concerns with their clothing. In contrast, men who score similarly high on the Body Surveillance scale might not be very concerned with their clothing, but might score high on other items (e.g., Body Surveillance: "During the day, I think about how I look many times;" Appearance Evaluation: "I like my looks just the way they are"). The overall mean scores would give the impression that men and women have similar experiences of body surveillance, but analysis of measurement variance would show that different items are contributing to these similarly high scores.

Importantly, however, the ability to compare means across groups and draw valid conclusions assumes that the measures used to assess the construct of interest operate in the same way across groups such that a given scale score reflects similar levels of the latent construct for each group. If measures do not meet this assumption (or if the assumption has not been rigorously tested), any observed group differences become difficult to interpret and researchers are unable to determine whether such findings reflect true group differences, or are instead the product of measurement bias. As the samples used to develop and validate many body image measures have, presumably, often been comprised primarily of heterosexual White college women (e.g., Cash & Fleming, 2002; McKinley & Hyde, 1996), the possibility for measurement bias in more diverse groups is significant, and the need to evaluate such bias is imperative.

Measurement invariance testing is used to evaluate the assumption that scales operate similarly across groups (Byrne et al., 1989). This approach involves applying increasingly restrictive equivalence constraints on elements of the scale structure using multi-group confirmatory factor analysis (CFA) and examining changes in model fit at each step. Significant decrements in model fit at increasing levels of constraint indicate invariance, or lack of similarity in scale functioning across groups. More specifically, measurement invariance testing first seeks to evaluate *configural invariance* (also referred to as equal form invariance) to assess the equivalence of a scale's factor structure across groups. Next, the assumption of *metric invariance* (also referred to as equal factor loading or weak factorial invariance) is tested to assess equivalence in factor loadings across groups. This step seeks to evaluate whether identified latent factors have the same meaning across groups. Finally, *scalar invariance* (also referred to as equal intercepts or strong factorial invariance) is evaluated to assess the equality of indicator intercepts, or whether levels of the latent factors have the same meaning across groups. The establishment of scalar invariance indicates that observed group differences on latent scores are not attributable to measurement differences, and therefore group factor scores can be meaningfully compared (Brown, 2015).

Good-practice guidelines for scale development highlight the importance of establishing measurement invariance (Swami & Barron, 2019), which has led researchers to investigate such invariance in existing body image scales (e.g., Alcaraz-Ibáñez & Sicilia, 2018; Lemoine et al., 2018; Meneses et al., 2019; Sladek et al., 2018). The limited number of existing measurement invariance studies suggest that some body image scales are vulnerable to measurement variance across genders (Tod et al., 2012), races (Kelly et al., 2011), and weight statuses (Pakpour et al., 2019). These findings raise a major concern – can we trust that the many of the conclusions regarding group differences in existing research are not heavily biased by measurement variance across groups?

A scale's measurement properties should be established as equivalent across demographic groups if the scale is intended for use in heterogenous samples with respect to these groups (Brown, 2015). Therefore, the present study addressed this concern by examining measurement invariance across five key demographic variables (i.e., age group, gender, sexual orientation, race, and weight status) for five commonly used measures: the Appearance Evaluation and Overweight Preoccupation subscales of the Multidimensional Body-Self Relations Questionnaire (Brown et al., 1990; Cash, 2000), Body Image Quality of Life Inventory (Cash & Fleming, 2002), Body Surveillance subscale of the Objectified Body Consciousness Scale (McKinley & Hyde, 1996), and Sociocultural Attitudes Towards Appearance Questionnaire-4 (Schaefer et al., 2015).

Without establishing measurement invariance across groups, differences in scale means across groups could potentially reflect measurement bias rather than true group differences (Brown, 2015). Therefore, as body image researchers may wish to examine differences in scores on commonly used body image measures across groups defined by age, gender, sexual orientation, race, and weight status, measurement invariance testing across such groups is crucial for drawing meaningful conclusions. The current study significantly extends the extant literature base, providing the most comprehensive evaluation of measurement invariance across key demographic variables within multiple commonly used measures of body image, and represents the first paper to examine measurement invariance within any measure of body image across sexual orientations. Results from the current study will aid in the interpretation of findings regarding body image experiences among different groups.

2. Method

2.1. Participants

Data were drawn from The U.S. Body Project I, described in the Procedure section. The sample was restricted to include only participants who completed the full survey and who fit the following criteria: (a) reported currently living in the United States; (b) completed all key body image items; (c) were aged 18–65; (d) had body mass indexes (BMI) ranging from 14.50 to 50.50 kg/m^2 based on self-reported height and weight. Age and BMI restrictions were placed on the sample to prevent outliers or mis-entered values from having undue influence on the effect size estimates. A total of 13,518 people clicked on the survey, 12,571 answered the first question, and 12,151 completed the full survey. After applying the inclusion criteria, this created the base dataset for The U.S. Body Project I of 11,620 participants. Key demographic characteristics of the sample are shown in Table 1. The sample was predominantly between the ages of 18–35 years (65.3%), female (54.4%), heterosexual (88.3%), White (75.2%), and classified as with BMIs in the 18.5-24.9 range (39.0%). For more detailed demographics and a discussion of how the current sample compares to nationally representative datasets, please see Frederick and Crerand et al. (2022).

2.2. Procedure and Overview of The U.S. Body Project I

The first author's university institutional review board approved the study. Adult participants were recruited via Amazon Mechanical Turk, a widely used online panel system used by researchers to access adult populations (Berinsky et al., 2012, Buhrmester et al., 2011, Kees et al., 2017; Paolacci et al., 2010; Robinson et al., 2019). Participants were paid 51 cents for taking the survey. The survey was advertised with the title "Personal Attitudes Survey" and the description explained that "We are measuring personal attitudes and beliefs. The survey will take roughly 10–15 minutes to complete." The general wording of the advertisement was used to avoid selectively recruiting people particularly interested in body image. After clicking on the advertisement, the participants read a consent form providing more details about the content of the study, including that it would contain items related to sex, love, work, and appearance. They were then given the option to continue with the survey or exit.

After providing informed consent, participants completed the numerical textbox questions (e.g., hours per week worked, number of times in love, sex frequency per week, longest relationship), followed by appearance evaluation (Cash, 2000), the SATAQ-4 (Schaefer et al., 2015), face satisfaction (Frederick & Kelly et al., 2016), overweight preoccupation (Cash, 2000), body image quality of life (Cash & Fleming, 2002), appearance surveillance (McKinley & Hyde, 1996), and finally demographics.

This manuscript is part of a series of papers emerging from The U.S. Body Project I. This project invited over twenty body image and eating disorder researchers, four sexuality researchers, and six computational scientists to apply their content and data-analytic expertise to the dataset. This project resulted in the following set of 12 papers for this special issue.

The first two papers examine how demographic factors (gender, sexual orientation, BMI, age, race) are related to body satisfaction and overweight preoccupation (Frederick, Crerand, et al., 2022) and to measures derived from objectification theory and the tripartite influence model, including body surveillance, thin-ideal and muscular/athletic ideal internalization, and perceived peer, family, and media pressures (Frederick, Pila, et al., 2022). The second set of papers examine how these measures and demographic factors predict sexuality-related body image (Frederick, Gordon, et al., 2022) and face satisfaction (Frederick, Reynolds, et al., 2022).

The third set of papers use structural equation modelling to examine the links between sociocultural appearance concerns and body satisfaction among women and across BMI groups (Frederick, Tylka, Rodgers, Pennesi, et al., 2022), among men and across different BMI groups (Frederick, Tylka, Rodgers, Convertino, et al., 2022), across racial groups (Frederick, Schaefer, et al., 2022) and across sexual orientations (Frederick, Hazzard, Schaefer, Rodgers, et al., 2022).

The fourth set of papers focus on measurement issues by examining measurement invariance of the scales across different demographic groups (current paper) and conducting a psychometric evaluation of an abbreviated version of the Body Image Quality of Life Inventory (Hazzard, Schaefer, Thompson, Murray, & Frederick, 2022). Finally, the fifth set of papers uses machine learning modelling to compare the effectiveness of nonlinear models versus linear regression for predicting body image outcomes (Liang et al., 2022) and to use unsupervised machine learning hierarchical cluster models to identify how aspects of body image cluster differently across participants in multidimensional space (Rosenfield et al., 2022).

2.3. Measures

2.3.1. Objectified Body Consciousness Scale - Body Surveillance Subscale —Participants completed the 8-item Surveillance subscale of the Objectified Body Consciousness Scale (OBCS-Surveillance; McKinley & Hyde, 1996), which assesses the extent to which people monitor how they appear to others (e.g., "During the day, I think about how I look many times"). Responses were recorded on a 7-point Likert agreement scale with response options ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). Items were averaged, with higher scores indicating greater levels of surveillance (α = .86).

2.3.2. Multidimensional Body-Self Relations Questionnaire - Appearance Evaluation Subscale—The 7-item Appearance Evaluation subscale of the Multidimensional Body-Self Relations Questionnaire (MBSRQ-Appearance Evaluation), was used to measure feelings of physical attractiveness and satisfaction with one's appearance (e.g., "I like my looks just the way they are") (Brown et al., 1990; Cash, 2000). Responses were recorded on a 5-point Likert agreement scale with response options ranging from 1 (*Definitely Disagree*) to 5 (*Definitely Agree*). Items were averaged, with higher scores indicating more positive evaluations of appearance (α = .93).

2.3.3. Multidimensional Body-Self Relations Questionnaire - Overweight Preoccupation Subscale—The 4-item Overweight Preoccupation subscale of the Multidimensional Body-Self Relations Questionnaire (MBSRQ-Overweight Preoccupation), was used to measure fat anxiety, weight vigilance, dieting, and eating restraint (Brown et al., 1990; Cash, 2000). Responses to the first three questions were recorded on a 5-point Likert agreement scale with response options ranging from 1 (*Definitely Disagree*) to 5 (*Definitely Agree*), while responses to the last question were recorded on a frequency scale with response options ranging from 1 *{Never)* to 5 (*Very Often*). Items were averaged, with higher scores indicating more preoccupation with weight (α = .80).

2.3.4. Sociocultural Attitudes Towards Appearance Questionnaire-4—The Sociocultural Attitudes Towards Appearance Questionnaire-4 (SATAQ-4; Schaefer et al., 2015), which contains five subscales assessing perceived appearance pressures from family, peers, and media, as well as internalization of the thin ideal and muscular ideal, was also administered. An example of a pressure item was "I feel pressure from the media to look in better shape." Responses were recorded on a 5-point Likert scale (1= *Definitely Disagree*; 5 = *Definitely Agree).*

Of the 22 items in the original questionnaire, 21 were assessed in the present study ("I want my body to look like it has little fat" was inadvertently omitted from the survey). Responses were recorded on a 5-point Likert agreement scale with response options ranging from 1 (*Definitely Disagree*) to 5 (*Definitely Agree*). Items were averaged for each subscale, with higher scores indicating greater levels of perceived pressures and internalization ($\alpha_{\text{Thin-ideal Internalization}}$ = .84; $\alpha_{\text{Muscular-Ideal Internalization}}$ = .92; $\alpha_{\text{Family Appearance Pressures}}$ =.91; $\alpha_{\text{Peer Appearance Pressures}}$ =.93; $\alpha_{\text{Media Appearance Pressures}}$ =.97).

2.3.5. Body Image Quality of Life Inventory—Participants also completed the 19-item Body Image Quality of Life Inventory (BIQLI; Cash & Fleming, 2002), which assesses participant's beliefs about how their bodies affect their lives. Participants indicated whether their feelings about their bodies had positive, negative, or no effects on various aspects of their lives (e.g., "My day-to-day emotions," "How confident I feel in my everyday life," and "How happy I feel in my everyday life"). Participants responded on a 7-point Likert-type scale ranging from 1 (*Very Negative Effect*) to 7 (*Very Positive Effect).* Items were averaged, with higher scores indicating more positive effects of one's body image on quality of life (α = .96).

2.3.6. Demographics—Participants self-reported their age, gender, race, sexual orientation, height in feet and inches, and weight in pounds. Age was categorized into the following three groups: (1) 18–35 years, (2) 36–55 years, and (3) 56–65 years representing young adulthood, middle adulthood, and older adulthood (Petry, 2002). Different research traditions and researchers from different countries have different norms for labelling these groups, with some distinguishing "ethnicity" from "race," some eliminating the term race and relying solely on "ethnicity," some relying solely on the term "race," some using a hybrid of "race/ethnicity," and some referring to these groupings as "racial identities." The authorship team had a diverse set of views on the appropriate terminology. For brevity, we primarily rely on the term "race" throughout this manuscript to refer to these different identities, with the recognition that in some research traditions appropriate terminology differs. Participants indicate which category they belonged to from a list of identities and they could check all that apply (e.g., Asian, White, Black, Hispanic, Pacific Islander). Turning to sex and gender, in more recent surveys, researchers have commonly used a two-step approach to assessing sex assigned at birth and then gender with a diversity of options. The item in current study, however, mixed together a stem asking for people's "gender" with responses that conventional views would deem to be "sex:" male or female. Thus, depending on the interpretation of the question, participants could have been reporting their gender identity or their sex assigned at birth. Given this discrepancy, we refer to the variable as gender throughout the manuscript, but recognize for transgender participants, it could be unclear whether they selected their gender identity or their sex assigned at birth.

Using the self-reported height and weight data, we calculated BMI. We then divided participants into the traditional BMI categories used by the Centers for Disease Control: "underweight" (below 18.5 kg/m^2), "normal weight" (18.5–24.9 kg/m^2), "overweight" (25–29.9 kg/m^2), and "obese" (30 and above kg/m^2). We hasten to add that these widely-used categories were chosen as a heuristic so that the BMI results could be compared to existing studies, and do not represent uniform endorsement of the categories by the entire authorship team in terms of semantic accuracy or as clear indicators of a person's health status (e.g., see Tomiyama et al., 2016). To avoid any stigmatizing effects of these labels, we instead label these BMI groups as Lowest (Underweight), Low (Normal), Medium (Overweight), and High (Obese) BMI groups from this point forward.

2.4. Data Analysis

Descriptive statistics and internal consistency were computed with SPSS 25. Overall confirmatory factor analyses (CFAs) and evaluation of measurement invariance via multi-group CFAs were conducted using Mplus 8.6. As the sample contained only participants that completed the full survey, there were no missing data. Estimation via robust maximum likelihood was used, as there were five or more response categories for items on each measure and, per significant Kolmogorov-Smirnov tests and visual analysis of histograms and Q-Q plots, data were non-normal but not severely non-normal (Rosellini & Brown, 2021).

Overall and multi-group CFAs were conducted separately for each body image measure. Adequacy of model fit was judged by the following fit indices: comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root-mean square residual (SRMR). Values ≥ .95 for CFI, ≤ .06 for RMSEA, and ≤ .08 for SRMR indicate good model fit (Hu & Bentler, 1999). Values of .90 or higher for CFI, up to .10 for RMSEA, and up to .10 for SRMR indicate acceptable but mediocre model fit (Bentler, 1990; Browne & Cudeck, 1993; Hu & Bentler, 1995; MacCallum et al., 1996; Schermelleh-Engel & Müller, 2003). Models were deemed to have acceptable fit if most fit indices examined (i.e., at least two out of three) suggested acceptable fit, given evidence that individual fit indices and their associated cutoffs vary as a function of numerous factors (e.g., sample size, degrees of freedom, factor loadings) (Brown, 2015).

Initial CFAs were conducted in the full sample for each body image measure. For models with unacceptable fit in initial CFAs, model respecification guided by modification indices > 10.00 and theory was carried out in one-half of the sample (n = 5,810) that was computer-selected using a random number seed. Model respecification was conducted until acceptable fit was achieved in that half of the sample. Respecified models were then tested in the other half of the sample. When models for all body image measures achieved acceptable fit in overall CFAs, each model underwent multi-group CFA in the full sample to assess configural, metric, and scalar measurement invariance across age group, gender, sexual orientation, race, and weight status. Measurement invariance by sexual orientation was examined within each gender because sexual orientation is defined not only by the gender(s) one is romantically or sexually attracted to, but also by one's own gender. Only groups comprised of respondents that self-identified with a single demographic group (e.g., White) and which met or came very close to meeting the suggested minimum sample size of 200 (Kelloway, 2015) without being collapsed into an "Other" group were included in multi-group CFAs. Thus, multi-group CFAs utilized the full sample for evaluation of measurement invariance by gender and weight status, but 157 participants were excluded for multi-group CFAs by sexual orientation of small samples in identities other than heterosexual, bisexual, or gay/lesbian, and 920 participants were excluded for multi-group CFAs by race because of small samples in identities other than Asian, Black, Hispanic, or White.

The first step in each multi-group CFA assessed configural invariance, in which all factor loadings and item intercepts were free to vary across groups. The second step assessed metric (i.e., weak) invariance, in which the factor loadings were constrained to be equal across groups. The third step assessed scalar (i.e., strong) invariance, in which factor loadings and item intercepts were constrained across groups. Nested models (i.e., metric compared to configural; scalar compared to metric) were compared using the changes in CFI, RMSEA, and SRMR, as well as the Satorra-Bentler scaled χ^2 difference test (Satorra & Bender, 2010). Because χ^2 is sensitive to sample size, changes in CFI, RMSEA, and SRMR were given stronger consideration than the χ difference test (Putnick& Bomstein, 2016). If changes between nested models indicating worse model fit with the addition of parameter constraints are greater than .010 for CFI, .015 for RMSEA, and .030 for SRMR for metric invariance or .015 for SRMR for scalar invariance, measurement invariance across groups is not supported (Chen, 2007; Cheung & Rensvold, 2002). Notably, the stepwise progression from configural to scalar invariance places increasing equality constraints across a family of parameters in the CFA model (e.g., all item intercepts assumed to be equal). Thus, significant decrements in model fit between these nested models indicates that at least one parameter is noninvariant (i.e., full invariance is not supported), but should not be interpreted to mean that all parameters are noninvariant. Fit diagnostics (e.g., modification indices) can be used to identify the source of misfit and test for partial measurement invariance. Partial invariance is supported when some but not all parameters are equivalent. Research suggests that scale means can be compared across groups if partial metric and partial scalar invariance is observed (Brown, 2015; Byrne et al., 1989).

3. Results

3.1. Overall Confirmatory Analysis

Results of overall CFAs are reported in Table 2. Initial CFAs in the full sample indicated acceptable model fit for the hypothesized 5-factor SATAQ-4, 1-factor MBSRQ-Appearance Evaluation, and 1-factor MBSRQ-Overweight Preoccupation models, but unacceptable model fit for the hypothesized 1-factor OBCS-Surveillance and 1-factor BIQLI models.

Respecification of the OBCS-Surveillance model was conducted in one-half of the sample. Inspection of modification indices suggested allowing the error variances for one pair of items to correlate (modification index = 832.45). These items were the only two positively worded items in the scale (i.e., not reverse-scored); thus, covariance between these items likely comes from the directionality of item wording (Brown, 2015). Model fit was acceptable after correlating these error variances (CFI = .927, RMSEA = .089 with 90% CI = .084–.094, SRMR = .038). The respecified OBCS-Surveillance model achieved acceptable model fit in the other half of the sample as well (see Table 2).

Respecification of the BIQLI model was also conducted in one-half of the sample. Inspection of modification indices suggested allowing the error variances for five pairs of items to correlate. Each of these modification indices was also theoretically supported, as covariance for each item pair was likely related to content overlap between items. Specifically, modification indices suggested shared error variances between (1) items referring to ability to control weight and ability to control eating behaviors (modification index = 1,917.29), (2) items referring to enjoyment of sex life and feelings of acceptability as a sexual partner (modification index = 1,333.53), (3) items referring to happiness and confidence (modification index = 1,295.75), (4) items referring to relationships with friends and relationships with family members (modification index = 1,067.38), and (5) items referring to general life satisfaction and day-to-day emotions (modification index = 1,014.63). Model fit was acceptable after correlating the error variances between each of these item pairs (CFI = .894, RMSEA = .086 with 90% CI = .084–088, SRMR = .058). The respecified BIQLI model achieved acceptable model fit in the other half of the sample as well (see Table 2).

3.2. Evaluation of Measurement Invariance

Measurement invariance results by age group, gender, sexual orientation, race, and weight status from multi-group CFAs are reported in Table 3 for the SATAQ-4, Table 4 for the respecified OBCS-Surveillance model, Table 5 for the MBSRQ-Appearance Evaluation subscale, Table 6 for the MBSRQ-Overweight Preoccupation subscale, and Table 7 for the respecified BIQLI model. Several ΔRMSEA values were above the .015 threshold, but in the majority of these instances, the ΔRMSEA corresponded to improved model fit with the addition of parameter constraints. Full configural and metric measurement invariance were supported for all five body image measures across age groups, gender groups, sexual orientation groups among both males and females, racial groups, and weight status groups. Full scalar measurement invariance was also supported in the majority of analyses, with some exceptions detailed below.

3.2.1. Partial Scalar Invariance for the Sociocultural Attitudes Towards Appearance Questionnaire-4—Full scalar invariance of the SATAQ-4 across weight status groups was not supported, as evidenced by a ΔCFI of .017 indicating worse model fit with the addition of item intercept constraints. Modification indices suggested relaxing the intercept constraints for the items referring to (1) family encouragement of reducing body fat, (2) peer encouragement of becoming thinner, (3) media pressure to improve appearance, and (4) media pressure to look in better shape. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .010).

Intercepts for the family and peer weight loss encouragement items were highest for participants categorized as High BMI, indicating that for participants with the same overall SATAQ-4 score, those categorized as High BMI tended to endorse these items somewhat more frequently than participants in other weight status groups. On the other hand, intercepts for the media appearance pressure items were highest for participants categorized as Lowest BMI, indicating that for participants with the same overall SATAQ-4 score, those categorized as Lowest BMI tended to endorse these items somewhat more frequently than participants in other weight status groups.

3.2.2. Partial Scalar Invariance for the Objectified Body Consciousness Scale - Body Surveillance Subscale—Full scalar invariance of the respecified OBCS-Surveillance model was not supported across age groups or genders, as evidenced by ΔCFIs of .011 for both analyses indicating worse model fit with the addition of item intercept constraints. When examining scalar invariance across genders, modification indices suggested relaxing the intercept constraint for the item referring to time spent thinking about how one's body feels versus how it looks. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .007). The intercept for this item was highest for men, indicating that for participants with the same overall OBCS-Surveillance subscale score, men tended to endorse thinking about how one's body looks more than how it feels somewhat more frequently than women.

3.2.3. Partial Scalar Invariance for the Multidimensional Body-Self Relations Questionnaire - Appearance Evaluation Subscale—Full scalar invariance of the MBSRQ-Appearance Evaluation subscale was not supported across genders or weight status groups, as evidenced by ΔCFIs of .017 and .013 in these analyses, respectively, indicating worse model fit with the addition of item intercept constraints. When examining scalar invariance across genders, modification indices suggested relaxing the intercept constraints for the items referring to (1) liking the way one's clothes fit and (2) liking the way one looks without clothes on. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .006). The intercepts for both of these items were highest for women, indicating that for participants with the same overall MBSRQ-Appearance Evaluation subscale score, women tended to endorse liking the way their clothes fit and the way they look without clothes on somewhat more frequently than men.

When examining scalar invariance across weight status groups, modification indices suggested relaxing the intercept constraint for the item referring to liking the way one's clothes fit. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .007). The intercept for this item was highest for participants categorized as Lowest BMI, indicating that for participants with the same overall MBSRQ-Appearance Evaluation subscale score, those categorized as Lowest BMI tended to endorse liking the way their clothes fit somewhat more frequently than participants in other weight status groups.

3.2.4. Partial Scalar Invariance for the Multidimensional Body-Self Relations Questionnaire - Overweight Preoccupation Subscale—Full scalar invariance of the MBSRQ-Overweight Preoccupation subscale across weight status groups was not supported, as evidenced by a ΔCFI of .082, a ΔRMSEA of .035, and a ΔSRMR of .032 each indicating worse model fit with the addition of item intercept constraints. Modification indices suggested relaxing the intercept constraints for the items referring to (1) being conscious of small changes in weight and (2) worrying about being or becoming fat. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .003, ΔRMSEA = .004, ΔSRMR = .000).

The intercept for the item referring to being conscious of small changes in weight was highest for participants categorized as Lowest BMI, indicating that for participants with the same overall MBSRQ-Overweight Preoccupation subscale score, those categorized as Lowest BMI tended to endorse being conscious of small changes in weight somewhat more frequently than participants in other weight status groups. The intercept for the item referring to worrying about being or becoming fat was highest for participants categorized as Low BMI, indicating that for participants with the same overall MBSRQ-Overweight Preoccupation subscale score, those categorized as Low BMI tended to endorse worrying about being or becoming fat somewhat more frequently than participants in other weight status groups. The intercepts for both of these items were lowest for participants categorized as High BMI, indicating that for participants with the same overall MBSRQ-Overweight Preoccupation subscale score, those categorized as High BMI tended to endorse these items somewhat less frequently than participants in other weight status groups.

3.2.5. Partial Scalar Invariance for the Body Image Quality of Life Inventory

—Full scalar invariance of the respecified BIQLI across weight status groups was not supported, as evidenced by a ΔCFI of .015 indicating worse model fit with the addition of item intercept constraints. Modification indices suggested relaxing the intercept constraints for the items referring to (1) ability to control weight and (2) ability to control eating behaviors. When these item intercepts were allowed to vary, results supported partial scalar invariance (ΔCFI = .009).

The intercepts for both of these items were highest for participants categorized as Lowest BMI and lowest for participants categorized as High BMI, indicating that for participants with the same overall BIQLI score, those categorized as Lowest BMI tended to report the most positive beliefs about how their bodies affect their ability to control their weight and eating behaviors, while those categorized as High BMI tended to report the most negative beliefs about how their bodies affect their ability to control their weight and eating behaviors.

4. Discussion

4.1. Summary of Key Findings

Numerous measures have been developed to assess the multidimensional construct of body image and its correlates. Such measures have typically undergone initial development and validation testing using samples predominantly composed of heterosexual White college women of “normal” weight, raising questions about the appropriateness of these measures in assessing body image experiences among other populations. The present study examined measurement invariance across five key demographic characteristics (i.e., age group, gender, race, sexual orientation, weight status) within five frequently used measures of body image constructs. Results supported full strong invariance across sexual orientation and racial groups for all body image measures examined, as well as at least partial strong invariance across age, gender, and weight status groups. Importantly, even in instances of partial strong invariance, these results indicate that scale scores can be compared across examined groups without major concerns that the observed differences are being produced by measurement variance.

Results from the current study must also be contextualized within the existing research base evaluating measurement invariance within the five examined scales. Previous work using the MBSRQ subscales has demonstrated support for measurement invariance across age and gender (Rusticus & Hubley, 2006) as well as across Black and White women (Kelly et al., 2012). Thus, the current study extends previous findings by demonstrating strong invariance across additional racial groups (i.e., Hispanic and Asian), sexual orientation groups, and weight status groups for the MBSRQ-Appearance Evaluation subscale and MBSRQ-Overweight Preoccupation subscale.

Previous work using the OBCS-Surveillance subscale has demonstrated support for measurement invariance across Black and White women (Kelly et al., 2012), as well as adolescents drawn from the community compared to those with an eating disorder (Dakanalis et al., 2017). While some work has supported strong measurement invariance across gender in adolescent boys and girls (Dakanalis et al., 2017), other work has not demonstrated support for strong measurement invariance across gender in undergraduate men and women (Chen & Russo, 2010). Therefore, the current study adds to this literature by providing support for measurement invariance across multiple demographic groups within an adult community sample.

To date, only one other study has examined measurement invariance within the BIQLI. Results from that study supported strong measurement invariance across age and gender (Rusticus et al., 2008). Therefore, the current study provides additional evidence of gender-based measurement invariance within the BIQLI, as well as support for invariance across sexual orientation, racial, and weight status groups.

Analyses of the SATAQ-4 demonstrated strong measurement invariance across examined groups. To our knowledge, only one prior study has examined measurement invariance within the SATAQ-4, with results suggesting possible measurement bias across Black and White college women (Burnette et al., 2020). In addition, prior work has demonstrated strong gender-based measurement invariance in earlier versions of the SATAQ (e.g., SATAQ-3; Wheeler et al., 2011).

Although the measures generally showed full or partial invariance across different groups, some instances of partial non-invariance are worth further consideration. For example, some instances of partial non-invariance on the Appearance Evaluation subscale centered around gender differences in items related to clothing. When considering men and women with the same level of overall appearance evaluation, women tended to more strongly endorse liking the way their clothes fit. Women’s fashion is far more variable and designed to be more form fitting than men’s fashion, women are encouraged to invest in their wardrobes to enhance their appearances, and women’s appearance is evaluated based on their clothing. This may lead clothing and comfort with clothing to become a more central aspect of women’s appearance evaluation and lead to increased confidence in their appearance based on their clothing style relative to other aspects of appearance evaluation. Alternatively, this pattern could just as easily reflect women who are choosing looser fitting clothing and less objectifying clothing liking the way their clothing fits, which leads to less surveillance of their appearance and greater overall appearance evaluation. Although men may also invest in clothing as a way to enhance their appearance, men's clothing choices tend to be driven by practicality and lack of concern about appearance to a greater extent than women's clothing choices are (Frith & Gleeson, 2004). Additionally, confidence in one's nude appearance (i.e., liking how one looks without clothes on) was higher among women than among men with equivalent levels of appearance evaluation. This may be a result of a woman's nude appearance being considered a more central element of her sexiness or attractiveness than it is for men, which then impacts her evaluation of her appearance (Fredrickson & Roberts, 1997). These findings may indicate value in assessing clothing-related appearance evaluation and nude appearance evaluation in greater detail in new scales, particularly in light of some past research identifying links between body image and clothing style preferences (Chattaraman & Rudd, 2006).

Across several of the measures, there were instances of partial non-invariance across BMI groups. Among participants with equivalent scale scores, those with lower BMIs tended to endorse the weight-based cognitions to a greater degree (Overweight Preoccupation), reported liking how their clothes fit them more (Appearance Evaluation), reported more positive beliefs about how their body image impacts their ability to control their weight and their eating behaviors (Body Image Quality of Life Inventory), and reported less pressure from family and peers to lose weight but more pressure from the media to improve their appearance and look in better shape (SATAQ-4). Overall, these findings point to the fact that the experiences of people with different BMIs might contribute differently to aspects of their body image, particularly as it relates to social pressures and cognitions about eating and weight, and comfort with how their clothes fit them.

The current study is the first to evaluate measurement invariance across age, gender, sexual orientation, and weight status groups within the updated scale, and contributes to the understanding of possible measurement invariance across racial groups. More generally, the current study contributes to a broader movement towards systematically testing for measurement invariance in widely used measures in the eating disorder literature (Burke et al., 2017; Compte et al., 2019; Klimek et al., 2021; Rand-Giovannetti et al., 2020), and in psychology more generally (Dong & Dumas, 2020; Emerson et al., 2017; Iurino & Saucier, 2020).

4.2. Limitations and Strengths

Limitations of the current study warrant acknowledgement. First, the sample was drawn from a non-clinical population of Mechanical Turk users living within the United States, and therefore results may not be generalizable to clinical populations, or individuals from other countries or cultures (for a more detailed discussion of how this sample compares to national samples, see Frederick, Crerand, et al., 2022). An additional concern is that, in retrospect, the payment structure for the survey could have been higher, especially in light of recent discussions that have encouraged researchers to pay Mechanical Turk users at a rate consistent with, or higher than, minimum wage when listing surveys (Silberman et al., 2018).

Furthermore, although the current study is strengthened by the large overall sample size and ability to examine measurement invariance across multiple distinct groups, the study was nonetheless unable to examine measurement invariance across additional demographic groups for which the sample sizes were inadequate. For example, the current study did not examine measurement invariance among individuals who identified as multiracial. Future work is needed to test measurement invariance among these important, but understudied groups (Brewster et al., 2019; Ivezaj et al., 2010).

Finally, while the current study supported strong measurement invariance in the SATAQ-4 across all examined groups, it is notable that a single item ("I want my body to look like it has little fat") from the Internalization: Thin/Low Body Fat subscale was not administered to respondents. Therefore, the current study provides strong support for the comparison of group means across four SATAQ-4 subscales (i.e., the Internalization: Muscular/Athletic and three Pressures subscales), and tentative support for the comparison of group means across the Internalization: Thin/Low Body Fat subscale. Future work using the SATAQ-4 may seek to confirm the invariance of the full thin ideal internalization subscale across diverse samples.

One strength of the study was that is contained a sizable number of racial minority participants. This allowed for the demonstration of measurement invariance across racial groups. It is important to note, however, that the definitions of these racial groups and identities are bound by specific cultural traditions, and that there can be substantial variations in the experiences of subgroups within these racial groupings. Furthermore, there is likely important cross-cultural variation to consider. For example, the overweight preoccupation measure assesses reported behavioral attempts at weight change (e.g., dieting and fasting) and also cognitions about weight (e.g., constantly worrying about being or becoming fat). As a potential illustration, cultural values in Belize might emphasize that the body is not something that can easily be modified, breaking the link between cognitions and behavior. Additionally, women might engage in behaviors to change their weight because they work in the tourism industry catering to Western tourists, without experiencing negative cognitions or distress about their weight (Anderson-Fye, 2004).

Identified limitations are balanced by several notable strengths. First, the current study takes advantage of a large diverse community dataset collected from men and women across all 50 states. Therefore, while comparison groups were selected to represent specific demographic groups (e.g., males), the individuals within each group nonetheless represent a diverse cross-section of the United States, increasing confidence in the generalizability of results to other community-based samples in the United States. Furthermore, the study examined measurement invariance across five frequently used body image measures across demographic statuses that have been the focus of comparison in the body image literature. In doing so, our study was able to provide empirical support for projects that compare scores in examined body image measures across the aforementioned demographic groups, and to increase confidence in those studies' conclusions.

4.3. Conclusions

In summary, results from the current study support strong measurement invariance across age, gender, sexual orientation, race, and weight status groups for the SATAQ-4, OBCS-Surveillance subscale, MBSRQ-Appearance Evaluation subscale, MBSRQ-Overweight

Preoccupation subscale, and BIQLI. Findings suggest that these measures assess the constructs of interest similarly across multiple demographic groups, providing researchers with increased confidence in the validity of results from studies which compare scale means across these groups.

Funding:

This research was supported by the National Institute of Mental Health (grant number T32 MH082761), the National Institute of General Medical Science (1P20GM134969-01A1), and the Kay Family Foundation Data Analytic Grant.

References

Alcaraz-Ibáñez M, & Sicilia A (2018). Psychometric evaluation and sex invariance of the Spanish version of the Body and Appearance Self-Conscious Emotions Scale. Body Image, 25, 78–84. 10.1016/j.bodyim.2018.02.006 [PubMed: 29494829]

Anderson-Fye EP (2004). A "Coca-Cola" shape: Cultural change, body image, and eating disorders in San Andres, Belize. Culture, Medicine and Psychiatry, 28(4), 561–595. 10.1007/s11013-004-1068-4 [PubMed: 15847054]

Austin SB, Haines J, & Veugelers PJ (2009). Body satisfaction and body weight: Gender differences and sociodemographic determinants. BMC Public Health, 9, 313. 10.1186/1471-2458-9-313 [PubMed: 19712443]

Bentler PM (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107, 238–246. 10.1037/0033-2909.107.2.238 [PubMed: 2320703]

Berinsky AJ, Huber GA, & Lenz GS (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. Political Analysis, 20, 351–368. 10.1093/pan/mpr057

Brewster ME, Velez BL, Breslow AS, & Geiger EF (2019). Unpacking body image concerns and disordered eating for transgender women: The roles of sexual objectification and minority stress. Journal of Counseling Psychology, 66, 131–142. 10.1037/cou0000333 [PubMed: 30702325]

Brown TA (2015). Confirmatory factor analysis for applied research (2nd ed.). New York, NY: Guilford Press.

Brown TA, Cash TF, & Mikulka PJ (1990). Attitudinal body image assessment: Factor analysis of the Body-Self Relations Questionnaire. Journal of Personality Assessment, 55, 135–144. 10.1080/00223891.1990.9674053 [PubMed: 2231236]

Browne MW, & Cudeck R (1993). Alternative ways of assessing model fit. In Bollen KA & Long JS (Eds.), Testing Structural Equation Models. Newbury Park, CA: Sage.

Buhrmester M, Kwang T, & Gosling SD (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? Perspectives on Psychological Science, 6, 3–5. 10.1177/1745691610393980 [PubMed: 26162106]

Burke NL, Tanofsky-Kraff M, Crosby R, Mehari RD, Marwitz SE, Broadney MM, Shomaker LB, Kelly NR, Schvey NA, Cassidy O, Yanovski SZ, & Yanovski JA (2017). Measurement invariance of the Eating Disorder Examination in black and white children and adolescents. International Journal of Eating Disorders, 50(1), 758–768. 10.1002/eat.22713 [PubMed: 28370435]

Burnette CB, Boutté RL, Sosnowski DW, & Mazzeo SE (2020). Measurement invariance of the SATAQ-4 in Black and White undergraduate women. Body Image, 32, 85–93. 10.1016/j.bodyim.2019.11.004 [PubMed: 31846907]

Byrne BM, Shavelson RJ, & Muthén B (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 105, 456–466. 10.1037/0033-2909.105.3.456

Cash TF (2000). The Multidimensional Body-Self Relations Questionnaire user's manual. Available from the author at www.body-images.com

Author Manuscript
Author Manuscript
Author Manuscript

Cash TF, & Fleming EC (2002). The impact of body-image experiences: Development of the Body Image Quality of Life Inventory. International Journal of Eating Disorders, 31, 455–460. 10.1002/eat.10033 [PubMed: 11948650]

Cash TF, & Smolak L (2011). Body image: A handbook of science, practice, and prevention, 2nd ed. New York, NY: Guilford Press.

Chattaraman V, & Rudd NA (2006). Preferences for aesthetic attributes in clothing as a function of body image, body cathexis and body size. Clothing and Textiles Research Journal, 24(1), 46–61. 10.1177/0887302X0602400104

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling, 14, 464–504. 10.1080/10705510701301834

Chen FF, & Russo NF (2010). Measurement invariance and the role of body consciousness in depressive symptoms. Psychology of Women Quarterly, 34(3), 405–417. 10.1111/j.1471-6402.2010.01585.X

Cheung GW, & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233–255. 10.1207/S15328007SEM0902_5

Compte EJ, Nagata JM, Sepúlveda AR, Schweiger S, Sbdar LS, Silva BC, Bressan M, Rivas A, Menga S, Cortes C, Bidacovich G, López P, Muiños R, Rutsztein G, Torrente F, & Murray SB (2019). Confirmatory factor analysis and measurement invariance of the eating disorders examination-questionnaire across four male samples in Argentina. International Journal of Eating Disorders, 52(6), 740–745. 10.1002/eat.23075 [PubMed: 30912863]

Copen CE, Chandra A, & Febo-Vazquez F (2016). Sexual behavior, sexual attraction, and sexual orientation among adults aged 18-44 in the United States: Data From the 2011-2013 National Survey of Family Growth. National Health Statistics Reports, 88, 1–14.

Dakanalis A, Timko AC, Clerici M, Riva G, & Carrà G (2017). Objectified body consciousness (OBC) in eating psychopathology: Construct validity, reliability, and measurement invariance of the 24-item OBC Scale in clinical and nonclinical adolescent samples. Assessment, 24, 252–274. https://doi.org/10.1177/1073191115602553 https://doi.org/10.1016/j.cpr.2011.07.011 [PubMed: 26336907]

Dong Y, & Dumas D (2020). Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. Personality and Individual Differences, 160, 109956. 10.1016/j.paid.2020.109956

Emerson SD, Guhn M, & Gadermann AM (2017). Measurement invariance of the Satisfaction with Life Scale: Reviewing three decades of research. Quality of Life Research, 26(9), 2251–2264. 10.1007/s11136-017-1552-2 [PubMed: 28324322]

Fallon EA, Harris BS, & Johnson P (2014). Prevalence of body dissatisfaction among a United States adult sample. Eating Behaviors, 15, 151–158. 10.1016/j.eatbeh.2013.11.007 [PubMed: 24411768]

Frederick DA, Crerand CE, Brown TA, Perez M, Best CR, Brady JP, Compte EJ, Convertino AD, Cook-Cottone CP, Gordon AR, Malcarne VL, Nagata JM, Parent MC, Pila E, Pennesi J-L, Reynolds TA, Rodgers RF, Schaefer LM, Thompson JK, Tylka TL, & Murray SB (2022). Demographic predictors of body image satisfaction: The U.S. Body Project I. Body Image.

Forbes GB, & Frederick DA (2008). The UCLA Body Project II: Breast and body dissatisfaction among African, Asian, European, and Hispanic American college women. Sex Roles, 58, 449–457. 10.1007/s11199-007-9362-6

Frederick DA, & Essayli JH (2016). Male body image: The roles of sexual orientation and body mass index across five national U.S. studies. Psychology of Men & Masculinity, 17, 336–351. 10.1037/men0000031

Frederick DA, Forbes GB, Grigorian K, & Jarcho JM (2007). The UCLA Body Project I: Gender and ethnic differences in self-objectification and body satisfaction among 2,206 undergraduates. Sex Roles, 57, 317–327. 10.1007/s11199-007-9362-6

Frederick DA, Garcia JR, Gesselman AN, Mark KP, Hatfield E, & Bohrnstedt G (2020). The Happy American Body 2.0: Predictors of affective body satisfaction in two U.S. national Internet panel surveys. Body Image, 32, 70–84. 10.1016/j.bodyim.2019.11.003 [PubMed: 31830668]

Frederick DA, Gordon AR, Cook-Cottone CP, Brady JP, Reynolds TA, Alley J, Garcia JR, Brown TA, Best CR, Brady JP, Crerand CE, Compte EJ, Convertino AD, Malcarne VL, Nagata JM, Parent

MC, Perez M, Pila E, Pennesi J-L, Rodgers RF, Schaefer LM, Thompson JK, Tylka TL, & Murray SB (2022). Demographic and sociocultural predictors of sexuality-related body image and sexual frequency: The U.S. Body Project I. Body Image.

Frederick DA, Hazzard VM, Schaefer LM, Rodgers RF, Gordon AR, Tylka TL, Pennesi J-L, Convertino AD, Parent MC, Brady JP, Brown TA, Brady JP, Crerand CE, Compte EJ, Cook-Cottone CP, Malcarne VL, Nagata JM, Perez M, Pila E, Thompson JK, & Murray SB (2022). Sexual orientation differences in pathways from sociocultural and objectification constructs to body satisfaction: The U.S. Body Project I. Body Image.

Frederick DA, Kelly MC, Latner JD, Sandhu G, & Tsong Y (2016). Appearance concerns among White and Asian American women: Sociocultural predictors of body, face, and eye satisfaction. Body Image, 16, 113–125. 10.1016/j.bodyim.2015.12.002 [PubMed: 26808353]

Frederick DA, Pila E, Malcarne VL, Compte EJ, Nagata JM, Best CR, Brady JP, Brown TA, Crerand CE, Convertino AD, Cook-Cottone CP, Gordon AR, Parent MC, Pennesi J-L, Perez M, Reynolds TA, Rodgers RF, Schaefer LM, Thompson JK, Tylka TL, & Murray SB (2022). Demographic predictors of objectification theory and tripartite influence model constructs: The U.S. Body Project I. Body Image.

Frederick DA, Reynolds TA, Barrera CA, Alley J, Garcia JR, & Murray SB (2022). Demographic and sociocultural predictors of face image satisfaction: The U.S. Body Project I. Body Image.

Frederick DA, Schaefer LM, Hazzard VM, Rodgers RF, Tylka TL, Ong LQ, Pennesi J-L, Convertino AD, Parent MC, Brady JP, Brown TA, Brady JP, Crerand CE, Compte EJ, Cook-Cottone CP, Gordon AR, Malcarne VL, Nagata JM, Perez M, Pila E, Thompson JK, & Murray SB (2022). Racial identity differences in pathways from sociocultural and objectification constructs to body satisfaction: The U.S. Body Project I. Body Image.

Frederick DA, Tylka TL, Rodgers RF, Convertino AD, Pennesi J-L, Parent MC, Brady JP, Brown TA, Brady JP, Crerand CE, Compte EJ, Cook-Cottone CP, Gordon AR, Malcarne VL, Nagata JM, Perez M, Pila E, Schaefer LM, Thompson JK, & Murray SB (2022). Pathways from sociocultural and objectification constructs to body satisfaction among men: The U.S. Body Project I. Body Image.

Frederick DA, Tylka TL, Rodgers RF, Pennesi J-L, Convertino AD, Parent MC, Brady JP, Brown TA, Brady JP, Crerand CE, Compte EJ, Cook-Cottone CP, Gordon AR, Malcarne VL, Nagata JM, Perez M, Pila E, Schaefer LM, Thompson JK, & Murray SB (2022). Pathways from sociocultural and objectification constructs to body satisfaction among women: The U.S. Body Project I. Body Image.

Fredrickson BL, & Roberts TA (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. Psychology of Women Quarterly, 21, 173–206. 10.1111/j.1471-6402.1997.tb00108.x

Frith H, & Gleeson K (2004). Clothing and embodiment: Men managing body image and appearance. Psychology of Men & Masculinity, 5(1), 40–48. 10.1037/1524-9220.5.1.40

Fryar CD, Kruszon-Moran D, Gu Q, & Ogden CL (2018). Mean body weight, height, waist circumference, and body mass index among adults: United States, 1999-2000 through 2015-2016. National Health Statistics Reports, 122, 1–16.

Grabe S, & Hyde JS (2006). Ethnicity and body dissatisfaction among women in the United States: A meta-analysis. Psychological Bulletin, 132, 622–640. 10.1037/0033-2909.132.4.622 [PubMed: 16822170]

Hazzard VM, Schaefer LM, Thompson JK, Rodgers RF, & Frederick DA (2022). An abbreviated 10-item, two-factor version of the Body Image Quality of Life Inventory (BIQLI-10): The U.S. Body Project I. Body Image.

He J, Sun S, Lin Z, & Fan X (2020). Body dissatisfaction and sexual orientations: A quantitative synthesis of 30 years research findings. Clinical Psychology Review, 101896. 10.1016/j.cpr.2020.101896 [PubMed: 32835956]

Hu L, & Bentler P (1995). Evaluating model fit. In Hoyle RH (Ed.), Structural Equation Modeling. Concepts, Issues, and Applications. London: Sage.

Hu LT, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55. 10.1080/10705519909540118

Iurino K, & Saucier G (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27 countries. Assessment, 27(2), 365–372. 10.1177/1073191118817916 [PubMed: 30596252]

Ivezaj V, Saules KK, Hoodin F, Alschuler K, Angelella NE, Collings AS, Saunders-Scott D, & Wiedemann AA (2010). The relationship between binge eating and weight status on depression, anxiety, and body image among a diverse college sample: A focus on bi/multiracial women. Eating Behaviors, 11, 18–24. 10.1016/j.eatbeh.2009.08.003 [PubMed: 19962116]

Kees J, Berry C, Burton S, & Sheehan K (2017). An analysis of data quality: Professional panels, student subject pools, and Amazon's Mechanical Turk. Journal of Advertising, 46, 141–155. 10.1080/00913367.2016.1269304

Kelloway EK (2015). Using Mplus for Structural Equation Modeling: A Researcher's Guide (2nd ed). Thousand Oaks, CA: SAGE Publications, Inc.

Kelly NR, Mitchell KS, Gow RW, Trace SE, Lydecker JA, Bair CE, & Mazzeo S (2012). An evaluation of the reliability and construct validity of eating disorder measures in white and black women. Psychological Assessment, 24, 608–617. 10.1037/a0026457 [PubMed: 22149327]

Klimek P, Convertino AD, Pennesi JL, Gonzales IV M, Roesch SC, Nagata JM, & Blashill AJ (2021). Confirmatory factor and measurement invariance analyses of the Eating Disorder Examination Questionnaire in sexual minority men and women. International Journal of Eating Disorders, 57(5), 745–754. 10.1002/eat.23488

Liang D, Frederick DA, Eiroa-Lledo E, Rosenfield NS, Berardi V, Linstead E, Maoz U (2022). Examining the utility of nonlinear machine learning versus linear regression analyses for predicting body image outcomes: The U.S. Body Project I. Body Image.

Lemoine JE, Konradsen H, Jensen AL, Roland-Lévy C, Ny P, Khalaf A, & Torres S (2018). Factor structure and psychometric properties of the Body Appreciation Scale-2 among adolescents and young adults in Danish, Portuguese, and Swedish. Body Image, 26, 1–9. 10.1016/j.bodyim.2018.04.004 [PubMed: 29772464]

MacCallum RC, Browne MW, & Sugawara HM (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1, 130–149. 10.1037//1082-989X.1.2.130

McKinley NM, & Hyde JS (1996). The Objectified Body Consciousness Scale: Development and validation. Psychology of Women Quarterly, 20, 181–215. 10.1111/j.1471-6402.1996.tb00467.x

Meneses L, Torres S, Miller KM, & Barbosa MR (2019). Extending the use of the Body Appreciation Scale-2 in older adults: A Portuguese validation study. Body Image, 29, 74–81. 10.1016/j.bodyim.2019.02.011 [PubMed: 30861485]

Morrison MA, Morrison TG, & Sager CL (2004). Does body satisfaction differ between gay men and lesbian women and heterosexual men and women? A meta-analytic review. Body Image, 1, 127–138. 10.1016/j.bodyim.2004.01.002 [PubMed: 18089146]

Murray SB, Nagata JM, Griffiths S, Calzo JP, Brown TA, Mitchison D, … & Mond JM (2017). The enigma of male eating disorders: A critical review and synthesis. Clinical Psychology Review, 57, 1–11. 10.1016/j.cpr.2017.08.001 [PubMed: 28800416]

Pakpour AH, Tsai M, Lin Y, Strong C, Latner., Fung XCC, Lin CY, & Tsang HWH (2019). Psychometric properties and measurement invariance of the Weight Self-Stigma Questionnaire and Weight Bias Internalization Scale in children and adolescents. International Journal of Clinical and Health Psychology, 19, 150–159. 10.1016/j.ijchp.2019.03.001 [PubMed: 31193103]

Paolacci G, Chandler J, & Ipeirotis PG (2010). Running experiments on Amazon Mechanical Turk. Judgment and Decision Making, 5, 411–419. 10.2139/ssrn.1626226

Peplau LA, Frederick DA, Yee C, Maisel N, Lever J, & Ghavami N (2009). Body image satisfaction in heterosexual, gay, and lesbian adults. Archives of Sexual Behavior, 38, 713–725. 10.1007/s10508-008-9378-1 [PubMed: 18712469]

Putnick DL, & Bornstein MH (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review, 41, 71–90. 10.1016/j.dr.2016.06.004 [PubMed: 27942093]

Rand-Giovannetti D, Cicero DC, Mond JM, & Latner JD (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. 10.1177/1073191117738046 [PubMed: 29094603]

Rosenfield NS, Eiroa-Lledo E, Frederick DA, Liang D, Berardi V, Maoz U, & Linstead E (2022). Clusters of body image concerns identified through unsupervised learning hierarchical clustering algorithms: The U.S. Body Project I. Unpublished manuscript.

Roberts A, Cash TF, Feingold A, & Johnson BT (2006). Are black-white differences in females' body dissatisfaction decreasing? A meta-analytic review. Journal of Consulting and Clinical Psychology, 74, 1121–1131. 10.1037/0022-006X.74.6.1121 [PubMed: 17154741]

Robinson J, Rosenzweig C, Moss AJ, & Litman L (2019). Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool. Plos One, 14. 10.1371/journal.pone.0226394

Rusticus SA, & Hubley AM (2006). Measurement invariance of the Multidimensional Body-Self Relations Questionnaire: Can we compare across age and gender? Sex Roles: A Journal of Research, 55(11–12), 827–842. 10.1007/s11199-006-9135-7

Rusticus SA, Hubley AM, & Zumbo BD (2008). Measurement invariance of the Appearance Schemas Inventory--Revised and the Body Image Quality of Life Inventory across age and gender. Assessment, 15, 60–71. 10.1177/1073191107306805 [PubMed: 18258732]

Satorra A, & Bender PM (2010). Ensuring positiveness of the scaled chi-square test statistic. Psychometrika, 75(2), 243–248. [PubMed: 20640194]

Schaefer LM, Burke NL, Calogero RM, Menzel JE, Krawczyk R, & Thompson JK (2018). Self-objectification, body shame, and disordered eating: Testing a core mediational model of objectification theory among White, Black, and Hispanic women. Body Image, 24, 5–12. 10.1016/j.bodyim.2017.10.005 [PubMed: 29172061]

Schaefer LM, Burke NL, Thompson JK, Dedrick RF, Heinberg LJ, Calogero RM, Bardone-Cone AM, Higgins MK, Frederick DA, Kelly M, Anderson DA, Schaumberg K, Nerini A, Stefanile C, Dittmar H, Clark E, Adams Z, Macwana S, Klump KL, Vercellone AC, Paxton SJ, & Swami V (2015). Development and validation of the Sociocultural Attitudes Towards Appearance Questionnaire-4 (SATAQ-4). Psychological Assessment, 27, 54–67. 10.1037/a0037917 [PubMed: 25285718]

Schermelleh-Engel K, & Müller H (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Research, 8, 28–74. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.509.4258&rep=rep1&type=pdf

Shriver LH, Harrist AW, Page M, Hubbs-Tait L, Moulton M, & Topham G (2013). Differences in body esteem by weight status, gender, and physical activity among young elementary school-aged children. Body Image, 10, 78–84. 10.1016/j.bodyim.2012.10.005 [PubMed: 23228485]

Silberman MS, Tomlinson B, LaPlante R, Ross J, Irani L, & Zaldivar A (2018). Responsible research with crowds: Pay crowdworkers at least minimum wage. Communications of the ACM, 61, 39–41. 10.1145/3180492

Sladek MR, Salk RH, & Engeln R (2018). Negative body talk measures for Asian, Latina(o), and White women and men: Measurement equivalence and associations with ethnic-racial identity. Body Image, 25, 66–77. 10.1016/j.bodyim.2018.02.005 [PubMed: 29486394]

Swami V, & Barron D (2019). Translation and validation of body image instruments: Challenges, good practice guidelines, and reporting recommendations for test adaptation. Body Image, 31, 204–220. 10.1016/j.bodyim.2018.08.014 [PubMed: 30220631]

Tiggemann M (2004). Body image across the adult life span: Stability and change. Body Image, 1, 29–41. 10.1016/S1740-1445(03)00002-0 [PubMed: 18089139]

Tod D, Hall G, & Edwards C (2012). Gender invariance and correlates of the Drive for Leanness Scale. Body Image, 9, 555–558. 10.1016/j.bodyim.2012.06.004 [PubMed: 22795653]

Tomiyama AJ, Hunger JM, Nguyen-Cuu J, & Wells C (2016). Misclassification of cardiometabolic health when using body mass index categories in NHANES 2005–2012. International Journal of Obesity, 40, 883–886. 10.1038/ijo201617 [PubMed: 26841729]

Wheeler DL, Vassar M, & Hale WD (2011). A gender-based measurement invariance study of the Sociocultural Attitudes Toward Appearance Questionnaire-3. Body Image, 8, 168–172. 10.1016/j.bodyim.2010.12.004 [PubMed: 21354879]

Wilton LS, Rattan A, & Sanchez DT (2018). White's perceptions of biracial individuals' race shift when biracials speak out against bias. Social Psychological and Personality Science, 9, 953–961. 10.1177/1948550617731497

Wilton LS, Sanchez DT, & Garcia JA (2013). The stigma of privilege: Racial identity and stigma consciousness among biracial individuals. Race and Social Problems, 5, 41–56. 10.1007/s12552-012-9083-5

Highlights

- Measurement invariance of commonly used body image scales is largely unknown.
- We tested invariance by age, gender, race, sexual orientation, and weight status.
- Measures tested include Appearance Evaluation and Overweight Preoccupation.
- Additional measures tested included the SATAQ-4, BIQLI, and Body Surveillance.
- Strong (scalar) measurement invariance was supported across measures and groups.

Table 1.

Sample Characteristics

	Mean (*SD*)
Age (years)	**34.12 (10.71)**
	Percent (*N*)
Age group	
18–35 years	65.3 (7,590)
36–55 years	28.8 (3,341)
56–65 years	5.9 (689)
Gender	
Male	45.6 (5,293)
Female	54.4 (6,327)
Sexual orientation	
Among males:	
Heterosexual	92.0 (4,869)
Bisexual	3.7 (194)
Gay/lesbian	3.7 (194)
Other	0.7 (36)
Among females:	
Heterosexual	85.3 (5,395)
Bisexual	9.5 (598)
Gay/lesbian	3.4 (213)
Other	1.9 (121)
Race	
White	75.2 (8,742)
Black	6.7 (774)
Hispanic	4.0 (470)
Asian	6.1 (714)
Other	7.9 (920)
Weight status	
Lowest BMI (“Underweight”)	1.6 (190)
Low BMI (“Normal weight”)	39.0 (4,535)
Medium BMI (“Overweight”)	31.3 (3,632)
Highest BMI (“Obese”)	28.1 (3,263)

Note. SD = standard deviation; BMI = body mass index

Table 2.

Confirmatory Factor Analysis Fit Indices for Measures Assessing Body Image Constructs

Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR
SATAQ-4	15,364.41 (179) ***	.897	.085 (.084, .087)	.051
OBCS-Surveillance	3,727.96 (20) ***	.849	.126 (.123, .130)	.061
Respecified OBCS-Surveillance	949.97 (19) ***	.925	.092 (.087, .097)	.039
MBSRQ-Appearance Evaluation	1,980.45 (14) ***	.946	.110 (.106, .114)	.031
MBSRQ-Overweight Preoccupation	356.90 (2) ***	.968	.124 (.113, .135)	.031
BIQLI	27,591.96 (152) ***	.771	.125 (.123, .126)	.069
Respecified BIQLI	6,280.86 (147) ***	.900	.085 (.083, .087)	.053

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; SATAQ-4 = Sociocultural Attitudes Towards Appearance Questionnaire-4; OBCS = Objectified Body Consciousness Scale; MBSRQ = Multidimensional Body-Self Relations Questionnaire; BIQLI = Body Image Quality of Life Inventory. Original models tested in full sample (N = 11,620); respecified models tested in cross-validation sample (N = 5,810).

* $p < .05$,

** $p < .01$,

*** $p < .001$.

Table 3.

Measurement Invariance for the Sociocultural Attitudes Towards Appearance Questionnaire-4

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δ*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Testing invariance by age group: 18–35 years (n = 7,590) vs. 36–55 years (n = 3,341) vs. 56–65 years (n = 689)									
Configural	15,585.36 (537)***	.896	.085 (.084, .086)	.052	--	--	--	--	Yes
Metric	15,957.65 (569)***	.894	.084 (.082, .085)	.053	272.33 (32)***	.002	.001	.001	Yes
Scalar	16,527.69 (601)***	.890	.083 (.082, .084)	.054	500.97 (32)***	.004	.001	.001	Yes
Testing invariance by gender: men (n = 5,293) vs. women (n = 6,327)									
Configural	15,403.71 (358)***	.896	.085 (.084, .086)	.051	--	--	--	--	Yes
Metric	15,722.28 (374)***	.894	.084 (.083, .085)	.052	251.30 (16)***	.002	.001	.001	Yes
Scalar	16,864.00 (390)***	.886	.085 (.084, .086)	.055	1,274.01 (16)***	.008	.001	.003	Yes
Testing invariance by sexual orientation among men: heterosexual (n = 4,869) vs. bisexual (n = 194) vs. gay (n = 194)									
Configural	9,338.09 (537)***	.864	.097 (.095, .098)	.062	--	--	--	--	Yes
Metric	9,486.82 (569)***	.863	.095 (.093, .096)	.063	84.77 (32)***	.001	.002	.001	Yes
Scalar	9,643.73 (601)***	.861	.093 (.091, .094)	.063	64.87 (32)***	.002	.002	.000	Yes
Testing invariance by sexual orientation among women: heterosexual (n = 5,395) vs. bisexual (n = 598) vs. lesbian (n = 213)									
Configural	7,609.79 (537)***	.917	.080 (.078, .081)	.041	--	--	--	--	Yes
Metric	7,735.50 (569)***	.915	.078 (.076, .080)	.042	54.61 (32)**	.002	.002	.001	Yes
Scalar	7,911.33 (601)***	.914	.077 (.075, .078)	.042	116.68 (32)***	.001	.001	.000	Yes
Testing invariance by race: White (n = 8,742) vs. Black (n = 774) vs. Hispanic (n = 470) vs. Asian (n = 714)									
Configural	15,642.66 (716)***	.897	.088 (.087, .089)	.051	--	--	--	--	Yes
Metric	15,838.79 (764)***	.896	.086 (.085, .087)	.051	68.85 (48)*	.001	.002	.000	Yes
Scalar	16,247.57 (812)***	.893	.084 (.083, .085)	.052	250.46 (48)***	.003	.002	.001	Yes
Testing invariance by weight status: Lowest BMI (n = 190) vs. Low BMI (n = 4,535) vs. Medium BMI (n = 3,632) vs. High BMI (n = 3,263)									

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Configural	16,056.72 (716)***	.899	.086 (.085, .087)	.053	--	--	--	--	Yes
Metric	16,668.19 (764)***	.895	.085 (.084, .086)	.056	498.31 (48)***	.004	.001	.003	Yes
Scalar	19,313.96 (812)***	.878	.089 (.087, .090)	.060	3,176.24 (48)***	.017	.004	.004	No

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; BMI = body mass index. Models were tested in full sample (N = 11,620).

* $p < .05$,

** $p < .01$,

*** $p < .001$.

Table 4.

Measurement Invariance for the Surveillance subscale of the Objectified Body Consciousness Scale

Invariance Model	$\chi^2(df)$	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Testing invariance by age group: 18–35 years (n = 7,590) vs. 36–55 years (n = 3,341) vs. 56–65 years (n = 689)									
Configural	1,961.68 (57)***	.921	.093 (.089, .096)	.040	--	--	--	--	Yes
Metric	2,046.43 (71)***	.919	.085 (.082, .088)	.044	51.92 (14)***	.002	.008	.004	Yes
Scalar	2,310.77 (85)***	.908	.082 (.079, .085)	.052	240.50 (14)***	.011	.003	.008	No
Testing invariance by gender: men (n = 5,293) vs. women (n = 6,327)									
Configural	1,850.44 (38)***	.924	.091 (.087, .094)	.039	--	--	--	--	Yes
Metric	1,988.87 (45)***	.919	.086 (.083, .089)	.047	106.36 (7)***	.005	.005	.008	Yes
Scalar	2,236.97 (52)***	.908	.085 (.082, .088)	.051	233.23 (7)***	.011	.001	.004	No
Testing invariance by sexual orientation among men: heterosexual (n = 4,869) vs. bisexual (n = 194) vs. gay (n = 194)									
Configural	857.04 (57)***	.921	.089 (.084, .095)	.041	--	--	--	--	Yes
Metric	904.60 (71)***	.918	.082 (.077, .087)	.044	24.16 (14)*	.003	.007	.003	Yes
Scalar	955.76 (85)***	.914	.076 (.072, .081)	.045	25.26 (14)*	.004	.006	.001	Yes
Testing invariance by sexual orientation among women: heterosexual (n = 5,395) vs. bisexual (n = 598) vs. lesbian (n = 213)									
Configural	1,096.08 (57)***	.923	.094 (.089, .099)	.039	--	--	--	--	Yes
Metric	1,143.96 (71)***	.921	.085 (.081, .090)	.041	11.63 (14)	.002	.009	.002	Yes
Scalar	1,201.08 (85)***	.917	.080 (.076, .084)	.041	24.23 (14)*	.004	.005	.000	Yes
Testing invariance by race: White (n = 8,742) vs. Black (n = 774) vs. Hispanic (n = 470) vs. Asian (n = 714)									
Configural	1,808.37 (76)***	.927	.092 (.089, .096)	.039	--	--	--	--	Yes
Metric	1,919.74 (97)***	.923	.084 (.081, .087)	.045	57.06 (21)***	.004	.008	.006	Yes
Scalar	2,110.46 (118)***	.916	.079 (.076, .082)	.052	141.48 (21)***	.007	.005	.007	Yes
Testing invariance by weight status: Lowest BMI (n = 190) vs. Low BMI (n = 4,535) vs. Medium BMI (n = 3,632) vs. High BMI (n = 3,263)									
Configural	1,891.41 (76)***	.926	.091 (.087, .094)	.040	--	--	--	--	Yes

Invariance Model	$\chi^2(df)$	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Metric	1,992.10 (97)***	.922	.082 (.079, .085)	.042	39.24 (21)**	.004	.009	.002	Yes
Scalar	2,209.64 (118)***	.914	.078 (.075, 081	.044	165.04 (21)***	.008	.004	.002	Yes

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; BMI = body mass index. Models were tested in full sample (*N*) 11,620).

* $p < .05$,

** $p < .01$,

*** $p < .001$.

Table 5.

Measurement Invariance for the Appearance Evaluation Subscale of the Multidimensional Body-Self Relations Questionnaire

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δ*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Testing invariance by age group: 18–35 years (n = 7,590) vs. 36–55 years (n = 3,341) vs. 56–65 years (n = 689)									
Configural	2,125.69 (42)***	.944	.113 (.109, .117)	.032	--	--	--	--	Yes
Metric	2,307.48 (54)***	.939	.104 (.100, .107)	.034	31.64 (12)**	.005	.009	.002	Yes
Scalar	2,621.16 (66)***	.931	.100 (.097, .103)	.038	259.98 (12)***	.008	.004	.004	Yes
Testing invariance by gender: men (n = 5,293) vs. women (n = 6,327)									
Configural	1,878.36 (28)***	.949	.107 (.103, .111)	.030	--	--	--	--	Yes
Metric	2,034.63 (34)***	.945	.101 (.097, .104)	.035	47.19 (6)***	.004	.006	.005	Yes
Scalar	2,637.18 (40)***	.928	.106 (.102, .109)	.047	677.94 (6)***	.017	.005	.012	No
Testing invariance by sexual orientation among men: heterosexual (n = 4,869) vs. bisexual (n = 194) vs. gay (n = 194)									
Configural	812.42 (42)***	.953	.102 (.096, .109)	.028	--	--	--	--	Yes
Metric	883.51 (54)***	.950	.094 (.088, .099)	.030	7.11 (12)	.003	.008	.002	Yes
Scalar	942.76 (94)***	.947	.087 (.082, .092)	.031	36.25 (12)***	.003	.007	.001	Yes
Testing invariance by sexual orientation among women: heterosexual (n = 5,395) vs. bisexual (n = 598) vs. lesbian (n = 213)									
Configural	1,128.06 (42)***	.946	.112 (.106, .117)	.032	--	--	--	--	Yes
Metric	1,262.66 (54)***	.940	.104 (.099, .109)	.037	33.71 (12)***	.006	.008	.005	Yes
Scalar	1,370.12 (66)***	.935	.098 (.093, .102)	.039	62.55 (12)***	.005	.006	.002	Yes
Testing invariance by race: White (n = 8,742) vs. Black (n = 774) vs. Hispanic (n = 470) vs. Asian (n = 714)									
Configural	1,876.19 (56)***	.946	.110 (.106, .115)	.032	--	--	--	--	Yes
Metric	2,061.61 (74)***	.941	.100 (.096, .104)	.036	36.69 (18)**	.005	.010	.004	Yes
Scalar	2,272.54 (92)***	.936	.094 (.091, .098)	.041	128.35 (18)***	.005	.006	.005	Yes
Testing invariance by weight status: Lowest BMI (n = 190) vs. Low BMI (n = 4,535) vs. Medium BMI (n = 3,632) vs. High BMI (n = 3,263)									
Configural	2,050.75 (56)***	.935	.111 (.107, .115)	.036	--	--	--	--	Yes

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δ*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Metric	2,354.32 (74)***	.926	.103 (.099, .107)	.050	181.18 (18)***	.009	.008	.014	Yes
Scalar	2,751.92 (92)***	.913	.100 (.097, .103)	.062	356.60 (18)***	.013	.003	.012	No

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; BMI = body mass index. Models were tested in full sample *N* = 11,620).

$^{*}p < .05$,

$^{**}p < .01$,

$^{***}p < .001$.

Table 6.

Measurement Invariance for the Overweight Preoccupation Subscale of the Multidimensional Body-Self Relations Questionnaire

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Testing invariance by age group: 18–35 years (n = 7,590) vs. 36–55 years (n = 3,341) vs. 56–65 years (n = 689)									
Configural	371.54 (6)***	.967	.125 (.115,.136)	.031	--	--	--	--	Yes
Metric	411.62 (12)***	.964	.093 (.085, .101)	.035	22.17 (6)**	.003	.032	.004	Yes
Scalar	486.78 (18)***	.958	.082 (.076, .088)	.037	65.52 (6)***	.006	.011	.002	Yes
Testing invariance by gender: men (n = 5,293) vs. women (n = 6,327)									
Configural	364.18 (4)***	.965	.124 (.114, .135)	.032	--	--	--	--	Yes
Metric	386.08 (7)***	.964	.097 (.088, .105)	.033	7.14 (3)	.001	.027	.001	Yes
Scalar	454.86 (10)***	.957	.088 (.081, .094)	.036	64.46 (3)***	.007	.009	.003	Yes
Testing invariance by sexual orientation among men: heterosexual (n = 4,869) vs. bisexual (n = 194) vs. gay (n = 194)									
Configural	172.88 (6)***	.963	.126 (.110, .143)	.030	--	--	--	--	Yes
Metric	187.65 (12)***	.961	.091 (.080, .103)	.034	9.27 (6)	.002	.035	.004	Yes
Scalar	194.43 (18)***	.961	.075 (.065, .084)	.034	4.86 (6)	.000	.016	.000	Yes
Testing invariance by sexual orientation among women: heterosexual (n = 5,395) vs. bisexual (n = 598) vs. lesbian (n = 213)									
Configural	211.79 (6)***	.964	.129 (.114, .144)	.034	--	--	--	--	Yes
Metric	230.08 (12)***	.962	.094 (.083, .105)	.034	2.68 (6)	.002	.035	.000	Yes
Scalar	247.12 (18)***	.960	.078 (.070, .087)	.036	10.42 (6)	.002	.016	.002	Yes
Testing invariance by race: White (n = 8,742) vs. Black (n = 774) vs. Hispanic (n = 470) vs. Asian (n = 714)									
Configural	345.15 (8)***	.967	.126 (.114, .137)	.031	--	--	--	--	Yes
Metric	381.43 (17)***	.964	.090 (.082, .097)	.035	20.02 (9)*	.003	.036	.004	Yes
Scalar	439.21 (26)***	.959	.077 (.071, .083)	.038	49.22 (9)***	.005	.013	.003	Yes
Testing invariance by weight status: Lowest BMI (n = 190) vs. Low BMI (n = 4,535) vs. Medium BMI (n = 3,632) vs. High BMI (n = 3,263)									
Configural	450.56 (8)***	.956	.138 (.127, .149)	.033	--	--	--	--	Yes
Metric	455.11 (17)***	.957	.094 (.087, .102)	.039	34.57 (9)***	.001	.044	.006	Yes
Scalar	1,291.95 (26)***	.875	.129 (.123, .136)	.071	939.16 (9)***	.082	.035	.032	No

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; BMI = body mass index. Models were tested in full sample (N = 11,620).

* $p < .05$,

** $p < .01$,

***$p < .001$.

Table 7.

Measurement Invariance for the Body Image Quality of Life Inventory

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δ*df*)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Testing invariance by age group: 18–35 years (n = 7,590) vs. 36–55 years (n = 3,341) vs. 56–65 years (n = 689)									
Configural	13,381.04 (441)***	.894	.087 (.086, .088)	.057	--	--	--	--	Yes
Metric	13,840.12 (477)***	.890	.085 (.084, .086)	.058	103.02 (36)***	.004	.002	.001	Yes
Scalar	14,461.85 (513)***	.885	.084 (.083, .085)	.058	397.18 (36)***	.005	.001	.000	Yes
Testing invariance by gender: men (n = 5,293) vs. women (n = 6,327)									
Configural	12,751.08 (294)***	.896	.085 (.084, .087)	.056	--	--	--	--	Yes
Metric	13,136.05 (312)***	.893	.084 (.083, .085)	.057	112.37 (18)***	.003	.001	.001	Yes
Scalar	13,791.08 (330)***	.888	.084 (.083, .085)	.059	596.93 (18)***	.005	.000	.002	Yes
Testing invariance by sexual orientation among men: heterosexual (n = 4,869) vs. bisexual (n = 194) vs. gay (n = 194)									
Configural	5,964.72 (441)***	.904	.085 (.083, .086)	.052	--	--	--	--	Yes
Metric	6,201.28 (477)***	.900	.083 (.081, .085)	.056	111.99 (36)***	.004	.002	.004	Yes
Scalar	6,369.63 (513)***	.898	.082 (.079, .082)	.057	44.03 (36)	.002	.001	.001	Yes
Testing invariance by sexual orientation among women: heterosexual (n = 5,395) vs. bisexual (n = 598) vs. lesbian (n = 213)									
Configural	9,118.61 (441)***	.883	.098 (.096, .099)	.062	--	--	--	--	Yes
Metric	9,387.03 (477)***	.880	.095 (.093, .097)	.063	61.35 (36)**	.003	.003	.001	Yes
Scalar	9,710.17 (513)***	.876	.093 (.091, .095)	.063	160.71 (36)***	.004	.002	.000	Yes
Testing invariance by race: White (n = 8,742) vs. Black (n = 774) vs. Hispanic (n = 470) vs. Asian (n = 714)									
Configural	12,642.96 (588)***	.894	.088 (.086, .089)	.057	--	--	--	--	Yes
Metric	13,086.78 (642)***	.891	.085 (.084, .086)	.059	88.22 (54)**	.003	.003	.002	Yes
Scalar	13,604.96 (696)***	.887	.083 (.082, .084)	.059	233.67 (54)***	.004	.002	.000	Yes
Testing invariance by weight status: Lowest BMI (n = 190) vs. Low BMI (n = 4,535) vs. Medium BMI (n = 3,632) vs. High BMI (n = 3,263)									

Invariance Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (Δdf)	ΔCFI	ΔRMSEA	ΔSRMR	Invariance?
Configural	13,318.24 (588)***	.894	.086 (.085, .088)	.058	--	--	--	--	Yes
Metric	13,829.47 (642)***	.890	.084 (.083, .085)	.060	118.10 (54)***	.004	.002	.002	Yes
Scalar	15,628.98 (696)***	.875	.086 (.085, .087)	.069	2,084.83 (54)***	.015	.002	.009	No

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root-mean square residual; BMI = body mass index. Models were tested in full sample ($N = 11,620$).

* $p < .05$,

** $p < .01$,

*** $p < .001$.