

# Eating Disorder Examination-Questionnaire: Evaluating factor structures and establishing measurement invariance with Asian/Hawaiian/Pacific Islander, Black, and White American college men

Neha J. Goel MS<sup>1</sup>  | Carolyn Blair Burnette PhD<sup>1,2</sup>  | Madison Weinstock BS<sup>1</sup> | Suzanne E. Mazzeo PhD<sup>1,3</sup>

<sup>1</sup>Department of Psychology, Virginia Commonwealth University, Richmond, Virginia, USA

<sup>2</sup>Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA

<sup>3</sup>Department of Pediatrics, Virginia Commonwealth University School of Medicine, Richmond, Virginia, USA

## Correspondence

Neha J. Goel, Department of Psychology, Virginia Commonwealth University, Richmond, VA 23284, USA.  
Email: goelnj@vcu.edu

## Funding information

National Heart, Lung, and Blood Institute, Grant/Award Number: T32HL150452; National Institute on Minority Health and Health Disparities, Grant/Award Number: F31 MD015679

Action Editor: Ruth Striegel Weissman

## Abstract

**Objective:** This study evaluated the factor structure and invariance of the Eating Disorder Examination-Questionnaire (EDE-Q) in a sample of Asian/Hawaiian/Pacific Islander ( $n = 163$ ), Black ( $n = 155$ ), and White ( $n = 367$ ) American university men.

**Method:** Twelve different EDE-Q factor structures reported in the literature were evaluated using multi-group confirmatory factor analyses, and measurement invariance assessed.

**Results:** A respecified four-factor structure proposed by Parker et al. (2016) showed superior fit and was invariant across groups. Significant differences emerged across all latent factors, with small to medium effects. Asian/Hawaiian/Pacific Islander men reported significantly higher scores on factors assessing Appearance Concern, Overvaluation of Shape/Weight, and Eating Concerns, and were more likely to endorse regular objective binge eating (OBE) and fasting episodes than their Black and White peers. Both White and Asian/Hawaiian/Pacific Islander men reported greater dietary restraint than Black men. Among this sample, frequencies of regular compensatory exercise ranged from 10% to 16%, fasting 6% to 14%, and OBEs 1% to 10%.

**Discussion:** Results provide further support for the use of alternate EDE-Q factor structures, especially among non-White men. In this study, Asian/Hawaiian/Pacific Islander men reported the highest levels of ED psychopathology relative to White and Black men, indicating they might be particularly vulnerable to EDs.

**Public Significance:** This study failed to find support for using the original Eating Disorder Examination-Questionnaire four-factor structure to detect disordered eating in Asian, Black, and White American college men. An alternate model proposed by Parker et al. in 2016 may be more appropriate. Asian men also reported the highest levels of eating psychopathology relative to their peers, suggesting they may be at high risk for developing eating disorders.

## KEYWORDS

African Americans, Asian Americans, college students, eating disorders, men, psychometrics

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *International Journal of Eating Disorders* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Eating disorders (EDs) are associated with high clinical impairment and poor quality-of-life (Schaumberg et al., 2017; Treasure et al., 2020), and are estimated to impact one in seven men in the United States (Ward et al., 2019). Notably, men and people of color are less likely to receive ED treatment and are historically underrepresented in measurement research, thereby exacerbating existing health disparities (Marques et al., 2011; Murray et al., 2017; Rand-Giovanetti et al., 2020).

The Eating Disorder Examination-Questionnaire (EDE-Q), a self-report measure of disordered eating cognitions, attitudes, and behaviors, is widely used in research and clinical settings (Fairburn & Beglin, 1994; Towne et al., 2017). The original EDE-Q comprises four factors and a global score. The four-factor model (i.e., Shape Concern, Weight Concern, Restraint, and Eating Concern) was developed based on qualitative interviews with patients with EDs, and extant etiological ED models, rather than an empirical measurement design process (Fairburn et al., 1993). The EDE-Q yields internally consistent and temporally stable global and subscale scores in American college students, although both of these scores are lower in samples of men, compared with women (Forbush et al., 2019; Rose et al., 2013). In community samples, the EDE-Q global score accurately distinguishes between individuals with and without EDs (Mond et al., 2004). Additionally, EDE-Q scores demonstrate high convergent validity with similar measures (e.g., Three Factor Eating Questionnaire; Stunkard & Messick, 1985; Bardone-Cone & Boyd, 2007; Berg et al., 2012).

Notably, the EDE-Q was initially normed and validated with a primarily White, Western, cisgender female sample (Berg et al., 2012; Fairburn & Beglin, 1994). Since its development, the EDE-Q has been translated into different languages (e.g., Italian, Spanish, Arabic; Calugi et al., 2017; Elder & Grilo, 2007; Melisse et al., 2021), and tested with samples living in Japan (Otani et al., 2021), Argentina (Compte et al., 2019), and Mexico (Penelo et al., 2013). Recent studies have also evaluated the utility of the EDE-Q in gender/sexual minority groups, and established norms for cisgender lesbian women (Nagata, Murray, Flentje, et al., 2020), cisgender bisexual women and men (Nagata, Compte, et al., 2020), and transgender women and men (Nagata, Murray, Compte, et al., 2020). Nonetheless, most of the data that established and continue to support the EDE-Q as a popular transdiagnostic ED assessment tool were collected from White female samples, raising questions about its generalizability and utility with other racial, ethnic, and gender groups (Berg et al., 2012). In addition to these concerns about the EDE-Q's generalizability, researchers have also questioned its four-factor structure (Barnes et al., 2012; Grilo et al., 2013; Hilbert et al., 2007; Rand-Giovanetti et al., 2020). Specifically, data suggest the Shape and Weight Concern subscales do not represent distinct constructs (Fairburn & Beglin, 1994; Hilbert et al., 2007; White et al., 2014). Several studies have proposed better-fitting alternatives, including a three-factor model (Peterson et al., 2007), a modified three-factor model (Grilo et al., 2013, 2015), and a unifactorial model (S. M. Byrne et al., 2010). However, most of these alternate models have been tested with primarily (or entirely) female samples (Rand-Giovanetti et al., 2020). Indeed, few studies have investigated potential gender differences in the EDE-Q factor structure

(i.e., Carey et al., 2019; Chan & Leung, 2015; Darcy et al., 2013; Grilo et al., 2015). All found that the factor structure differed in men and women; in particular, certain factor structures and items fit well for women, but not men.

Additional research regarding the EDE-Q's psychometrics in diverse samples of men is needed given the accumulating evidence that EDs affect individuals of all race/ethnicities, genders, and countries-of-origin (Cheng et al., 2019; Schaumberg et al., 2017). Extant research suggests that ED behaviors may manifest differently in men compared with other gender groups. For example, muscle dysmorphia (e.g., internalization of a muscular ideal, excessive exercise) is much more common in men than women (Mitchison & Mond, 2015; Murray et al., 2017). Unfortunately, the original EDE-Q items were designed to assess ED symptoms typically experienced by women (e.g., severe caloric restriction in pursuit of the thin ideal), making it difficult to determine whether the EDE-Q is sensitive enough to detect EDs in men. Thus, additional research investigating whether the EDE-Q offers utility as a viable screening tool for disordered eating in men is crucial.

Similarly, it is important to consider whether the EDE-Q yields valid scores in ethnically diverse populations. Serier et al. (2018) found the modified three-factor model proposed by Grilo et al. (2013) provided an acceptable fit in both Hispanic and non-Hispanic White women. However, the dietary restraint item loaded more strongly onto its latent factor for non-Hispanic White women, suggesting differences in the relevance and conceptualization of this item for Hispanic women (Serier et al., 2018). These differences might reflect variability in restraint between cultures, such that Latinx/Hispanic women are less likely to report this behavior, and when they do, it might be driven by different motivations than White women (e.g., need for control, rather than alteration of weight/shape; Marques et al., 2011; Perez & Warren, 2013).

Research has also identified significant differences between Black and White women on the EDE-Q. For example, scores above the clinical cutoff for Black women are less predictive of disordered eating behaviors, relative to White women (N. R. Kelly et al., 2012). N. R. Kelly et al. (2012) attributed this difference to the EDE-Q not accounting for factors uniquely affecting Black women that might influence eating behaviors, like racial discrimination and differing shape ideals (Boutté, 2020). Taken collectively, these findings suggest the EDE-Q might not adequately detect eating psychopathology in individuals from historically marginalized racial and ethnic groups (Burnette et al., 2020; N. R. Kelly et al., 2012; K. R. Kelly et al., 2012; Serier et al., 2018), which could perpetuate the marked disparities in ED detection and treatment (Goel et al., 2021; Goel, Jennings Mathis, et al., 2022; Hart et al., 2011; Marques et al., 2011).

The establishment of measurement invariance plays a crucial role in validating a measure, and having confidence in any group comparisons made regarding scores on it. Invariance indicates that the same underlying construct is being measured across groups, and that observed group differences are meaningful, and not the result of measurement error or item bias (B. M. Byrne et al., 2009; Chen, 2008; Putnick & Bornstein, 2016). Because men and marginalized racial/ethnic groups are underrepresented in the ED literature (Egbert et al., 2022; Goel, Jennings Mathis, et al., 2022; Murray et al., 2017), it is important not only to

identify the appropriate factor structure for the EDE-Q in men, but also to ensure it is measuring the same constructs across racial and ethnic groups.

Given these limitations, this study evaluated the factor structure of the EDE-Q, and assessed measurement invariance in a sample of Asian/Hawaiian/Pacific Islander, Black, and White American college men. Given the consistent lack of support for the four-factor structure (e.g., Rand-Giovanetti et al., 2020), we did not expect it to hold for this sample. However, given the considerable number of factor structures represented in the literature, we did not develop specific hypotheses about the factor structure in our sample.

## 2 | METHOD

### 2.1 | Procedure and participants

To obtain a sufficiently sized sample for multi-group confirmatory factor analysis (CFA), data from four larger studies were combined. All study procedures occurred at the same large southeastern public, urban university. In 2020–2021 this university's student population identified as 37.1% male, 44.6% White, 17.9% Black, and 13.5% Asian/Hawaiian/Pacific Islander (Virginia Commonwealth University, 2021). For all studies, informed consent was obtained and data were collected online through Research and Electronic Data Capture (REDCap; Harris et al., 2009). Ethics approval for research with human subjects was obtained from the university's institutional review board. Recruitment occurred through the psychology department-sponsored participant pool and eligible individuals received course credit for completion. The four studies were described as assessing: (1) body image, and attitudes and behaviors related to eating and exercise; (2) the relation of ethnic identity to body appreciation; (3) the influence of weight history on eating attitudes and behaviors; and (4) factors associated with perceptions of women. Some participants included in this study were also included in Burnette et al. (2019) and Burnette and Mazzeo (2020). Though all university students  $\geq 18$  years old were eligible for these four studies, inclusion in the current sample was restricted to individuals self-identifying as male and non-Latinx White, Black, or Asian/Hawaiian/Pacific Islander. We were unable to include men identifying as Latinx due to insufficient recruitment of this group ( $n = 47$ ). Participant lists across studies were reviewed, and 14 duplicate responses were removed.

See Table 1 for participant demographics by racial group and overall. Slightly over half of the sample identified as non-Latinx White (53.6%), as Black (22.6%), and as Asian/Hawaiian/Pacific Islander (23.8%).

## 2.2 | Measures

### 2.2.1 | Demographics

Participants self-reported their age, year-in-school, gender (including categorical options of male, female, other, and a text box for self-

identification), height and weight, and racial and ethnic identities (including categorical options for Asian, Black, Hawaiian, Latinx, Native American, Pacific Islander, White, Other).

### 2.2.2 | Eating psychopathology

This study used the 36-item EDE-Q (Fairburn & Beglin, 1994), a measure of disordered eating attitudes and behaviors over the past 28 days. Its ordinal items are rated from 0 = *No days or Not at all* to 6 = *Every day or Markedly*, and are averaged to calculate four subscale (i.e., Shape Concern, Weight Concern, Restraint, and Eating Concern) and global scores. Higher scores reflect greater eating pathology. Frequency of binge and loss-of-control (LOC) eating, purging behaviors, and compensatory exercise are also captured. There is considerable evidence of the EDE-Q's internal consistency and test-retest reliability in undergraduate men and women (Forbush et al., 2019; Lavender et al., 2010; Luce et al., 2008; Rose et al., 2013).

## 2.3 | Data analysis

Data were cleaned in SPSS 27.0 and exported to R for analyses (R Core Team, 2018). We examined data for outliers and normality. CFA and measurement invariance were assessed using the lavaan package (Rosseel, 2012). Missing data were minimal (0%–1.2% per item) and missing completely at random (MCAR),  $\chi^2(332) = 343.294$ ,  $p = .32$ . The lavaan package uses list-wise deletion. Despite the small amount of missing data, we opted to impute missing values using expectation maximization to leverage all available cases for analyses. We used the WebPower package (Zhang & Yuan, 2018) to estimate whether sample size was sufficient to detect a root mean square error of approximation (RMSEA) value reflective of adequate model fit ( $\leq .08$ ; Hu & Bentler, 1999) based on the  $df$  for each model given 80% power and  $\alpha = .05$ . There were no significant differences across racial groups in current body mass index (BMI), age, or year-in-school. Thus, we did not include covariates in analyses.

### 2.3.1 | Confirmatory factor analysis

Given that previous studies have yielded mixed findings regarding the EDE-Q's factor structure (Rand-Giovanetti et al., 2020), prior to assessing invariance, we conducted CFAs of the original four-factor structure (Fairburn & Beglin, 1994) and 11 alternate factor structures found in the literature to determine the best-fitting model. Those assessed included 2 four-factor models (Friborg et al., 2013; Parker et al., 2016), 6 three-factor structures (Darcy et al., 2013; Grilo et al., 2013; Hilbert et al., 2012; Peterson et al., 2007; White et al., 2014), a two-factor structure (Penelo et al., 2012), a one-factor structure including all subscale items (Pennings & Wojcieszowski, 2004), and a brief one-factor structure (Allen et al., 2011). Although several recent studies found the modified three-factor model yielded superior fit

TABLE 1 Descriptive statistics by race and overall

	AHP (n = 163)		Black (n = 155)		White (n = 367)		Overall (N = 685)	
	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI	M (SD)	95% CI
Age	19.44 (1.78)	[19.17, 19.71]	19.81 (2.87)	[19.35, 20.26]	19.90 (3.64)	[19.52, 20.27]	19.77 (3.12)	[19.54, 20.00]
BMI	24.30 (4.71)	[23.58, 25.02]	25.08 (5.71)	[24.18, 25.98]	24.15 (4.63)	[23.68, 24.62]	24.40 (4.92)	[24.03, 24.77]
Year in school	%(n)	95% CI	%(n)	95% CI	%(n)	95% CI	%(n)	95% CI
First	55.8% (91)	[.48, .63]	47.7% (74)	[.40, .56]	58.9% (216)	[.54, .64]	55.6% (381)	[.52, .59]
Second	18.4% (30)	[.12, .24]	27.1% (42)	[.20, .34]	18.3% (67)	[.14, .22]	20.3% (139)	[.17, .23]
Third	15.3% (25)	[.10, .21]	16.8% (26)	[.11, .23]	12.5% (46)	[.09, .16]	14.2% (97)	[.12, .17]
Fourth	9.82% (16)	[.05, .14]	8.39% (13)	[.04, .13]	8.7% (32)	[.06, .12]	8.9% (61)	[.07, .11]
Fifth+	-		-		1.6% (6)	[.003, .03]	0.8% (6)	[.001, .02]

Note: Descriptive statistics by race and overall. 95% CIs for year in school represent proportions rather than percentages. Overlap between the ranges of the lower and upper bounds of the CI across different groups is equivalent to a nonsignificant statistical test result.

Abbreviations: AHP, Asian/Hawaiian/Pacific Islander; CI, confidence interval.

(e.g., Grilo et al., 2013; Serier et al., 2018), it has only seven items, with just two indicators for two of the three factors. Similarly, the Parker et al. (2016) four-factor model has a factor with only two indicators. Because factors with fewer than three indicators often are unstable, unreliable, and can lead to under-identification of the model (a problem which is magnified when multiple factors have fewer than three indicators; Hair et al., 2010), we examined all model factor loadings and explored modification indices (MIs) of the next best-fitting model in the event that models with fewer than three indicators displayed the best fit. If theoretically justifiable modifications improved model fit to an acceptable level, we opted for a model with more indicators per latent factor to enhance confidence in the model's reliability and validity.

In all models, factors were correlated and estimated using weighted least squares mean and variance-adjusted estimation (WLSMV), which is appropriate for ordinal data (Kozioł & Bovaird, 2018; Li, 2016). We examined model fit using the  $\chi^2/df$  ratio ( $<3.0$ ), root mean square error of approximation (RMSEA  $\leq .08$ ), standardized root mean residual (SRMR  $\leq .08$ ), and the Comparative Fit and Tucker Lewis indices (CFI and TLI  $\geq .95$ ; Hu & Bentler, 1999). We also report the expected cross-validation index (ECVI) as another measure of absolute fit, but did not rely on it in model selection, in accordance with recommendations by Hu and Bentler (1998).

### 2.3.2 | Measurement invariance

We then assessed measurement invariance of the best-fitting model using guidelines provided by Wu and Estabrook (2016). Traditionally, measurement invariance is assessed using a series of nested CFA

models, where the factor structure is first constrained to be equal across groups (configural invariance), then factor loadings (metric invariance), and finally thresholds (scalar invariance). However, Wu and Estabrook (2016) have argued this method is not optimal for ordered categorical variables, and instead recommended constraining thresholds to equality prior to calculating factor loadings. Using the `measEq` function in `lavaan` (following guidelines provided by Svetina et al., 2020), we examined a series of nested CFA models, first constraining the factor structure (Model 1), then thresholds (Model 2), loadings (Model 3), and intercepts (Model 4) to equality. Each successive model retained the previous model's constraints. To compare model fit, we assessed changes in absolute fit indices (AFI) and conducted scaled chi-square difference tests using the `lavTestLRT` function. Guidelines for identifying non-invariance using  $\Delta AFI$  were derived based on models using continuous estimation methods (Chen, 2007). Using WLSMV estimation, there is evidence even small changes in AFI can indicate non-invariance, and that  $\Delta\chi^2$  might be more powerful than  $\Delta AFI$  (Kozioł & Bovaird, 2018; Sass et al., 2014). Thus, we evaluated invariance by consensus, choosing to probe for sources of non-invariance in the event of a significant  $\Delta\chi^2$  using the `lavTestScore` function.

Meaningful group comparisons can only be made on a measure with demonstrated invariance across groups. Thus, we evaluated group differences only in the event of an invariant model. To examine if EDE-Q scores differed across racial and ethnic groups, we constrained latent means to be equal (Model 5), with decrements in model fit indicating groups differed on subscale scores. We probed the nature of differences using one-way analyses of covariance (ANCOVAs) with Tukey's post hoc tests, adjusting for BMI. Partial  $\eta^2$  is presented as the effect size for overall ANCOVAs, with .01 = small, .06 = medium, and .14 = large effects, and Cohen's  $d$  is presented as

the effect size for pairwise comparisons of EDE-Q means, with .20 = small, .50 = medium, and .80 = large.

### 2.3.3 | Subscale scores and behavioral frequencies

Finally, we calculated subscale score means and standard deviations of the best fitting model by racial and ethnic group. We computed Spearman's correlations to assess internal consistency for two-item factors and McDonald's omega to assess internal consistency for factors with >2 indicators; values  $\geq .70$  demonstrate adequate reliability (Dunn et al., 2014; Nunnally, 1978). Additionally, we computed average frequencies of behavioral items and the proportions of men endorsing each behavioral item (with either any occurrence or regular occurrence). Because self-induced vomiting, laxative, and diuretic misuse are infrequently endorsed in non-clinical samples (e.g., Lavender et al., 2010), we collapsed these items into a category assessing "purging behaviors." Any occurrence of an ED behavior includes endorsement of the behavior  $\geq 1$  in the previous 28 days. Regular occurrence of objective binge episodes (OBEs), LOC eating, and purging behaviors was defined as  $\geq 4$  episodes in 28 days (Berg et al., 2012). Regular occurrence of compensatory exercise was defined as  $\geq 20$  episodes in 28 days, and fasting behaviors as going eight or more waking hours without eating  $\geq 13$  of 28 days (Lavender et al., 2010). We compared prevalence estimates of ED symptoms between groups using chi-square analyses. Number needed to take (NNT) is presented as the effect size, which represents the number of individuals in the reference group needed to sample to detect one more endorsement of ED symptoms in the comparison group (Kraemer et al., 2020). In each comparison, the group with the larger sample size was the reference group (e.g., Asian men were the reference group when rates were compared with Black men). NNT values are interpreted as  $>9$  = weak,  $4-9$  = moderate, and  $<4$  = strong.

## 3 | RESULTS

### 3.1 | Confirmatory factor analysis

We had adequate power across all models to detect an RMSEA  $\leq .08$  at 80% power given  $N = 685$  and  $\alpha = .05$ . Initially, the modified three-factor structure (Body Dissatisfaction, Overvaluation of Shape and Weight, and Restraint; Grilo et al., 2013, 2015) provided the best fit to the data, with fit indices in the acceptable ranges (see Table 2). However, as noted, this model has only seven items, with just two indicators for two of three latent factors, raising concerns about the measure's stability and reliability (Hair et al., 2010). Therefore, we examined MIs for the next best-fitting model, the Parker et al.'s (2016) four-factor model with Dietary Restraint, Eating Concern, Overvaluation of Shape/Weight, and Appearance Concern factors. MIs suggested allowing the error terms for two similarly worded items to correlate (#35: "How uncomfortable have you felt about seeing your body..." and #36: "How uncomfortable have you been about

others seeing your body") would improve model fit (MI: 115.58). Indeed, after respecifying the model, fit improved to acceptable ranges, scaled  $\Delta\chi^2(1) = 49.34$ ,  $p < .001$ . However, when evaluating model fit in each group, RMSEA and SRMR were above recommend thresholds in Black men. We examined MIs and standardized residuals (SRs) to evaluate areas of misfit (Brown, 2015). Although item #34 ("How concerned have you been about other people seeing you eat?") loaded highly onto its target factor, MIs suggested it might be cross-loading onto all other factors. Additionally, SRs suggested the model underestimated associations between #34 and items on the Appearance Concern and Overvaluation of Shape/Weight factors, but overestimated associations between #34 and items on the Restraint factor and other Eating Concern items. We evaluated MIs and SRs for the model in the overall sample, and observed the same pattern of results. Thus, we removed item #34 from the model and reevaluated model fit overall and within each group. Model fit improved substantially in all groups. Thus, we chose to proceed with the respecified Parker et al. (2016) model for invariance analyses. Although this model has one factor (Overvaluation of Shape/Weight) with only two indicators, the other factors have three to four indicators, providing better coverage for the constructs than the model by Grilo et al. (2013). Moreover, its four factors more closely mirror the original four factors captured by the original EDE-Q. See Figure 1 for factor loadings within each group.

### 3.2 | Measurement invariance

See Table 3 for an overview of invariance results. Fit indices were within acceptable ranges for the configural model (Hu & Bentler, 1999), providing support for the factor structure and its invariance across groups. The threshold invariance model (Model 2) also fit the data well. Both TLI and RMSEA improved, and CFI and SRMR remained unchanged relative to the configural model. The scaled chi-square difference test was not significant,  $\Delta\chi^2(88) = 100.71$ ,  $p = .17$ , providing support for the invariance of item thresholds. This suggests that item ratings reflected similar levels of the latent constructs across groups. Similarly, the loading invariance model (Model 3) fit the data well, with changes in fit indices comparable to the threshold invariance model. The nonsignificant scaled chi-square difference test provided further support for the invariance of factor loadings,  $(\Delta\chi^2[18] = 12.53$ ,  $p = .82)$ , suggesting each item contributed to its latent factor similarly across groups. Finally, there was no decrement in fit from the loadings to the intercepts model (Model 4), providing further support for invariance; RMSEA and TLI improved slightly, and the scaled chi-square difference test was not significant,  $\Delta\chi^2(18) = 18.99$ ,  $p = .39$ . Thus, any group differences in subscale scores should reflect *meaningfully different* levels of the latent construct.

With support for measurement invariance established, we assessed whether latent means differed between groups by constraining them to equality and comparing model fit (Model 5) to the intercepts model (Model 4). Although SRMR remained unchanged for the mean invariance model (Model 5), CFI ( $-.009$ ) and TLI decreased ( $-.007$ ) and RMSEA increased (.024). Furthermore, the scaled chi-



**TABLE 2** Fit statistics for CFA models of the Eating Disorder Examination-Questionnaire

Model	$\chi^2/df$	CFI	TLI	RMSEA	90% CI	SRMR	ECVI
Original four-factor structure (Fairburn & Beglin, 1994)	13.34	.911	.886	.134	[.130, .139]	.091	4.01
Four-factor (Friborg et al., 2013)	12.46	.907	.894	.129	[.125, .134]	.085	3.62
Four-factor (Parker et al., 2016)	7.44	.975	.968	.097	[.089, .105]	.052	0.74
Four-factor R1 (Parker et al., 2016)	5.02	.985	.980	.077	[.069, .085]	.050	0.58
<b>Four-factor R2 (Parker et al., 2016)</b>	<b>3.17</b>	<b>.993</b>	<b>.991</b>	<b>.056</b>	<b>[.047, .066]</b>	<b>.036</b>	<b>0.41</b>
Asian men	1.60	.992	.989	.061	[.036, .083]	.052	1.45
Black men	1.82	.986	.981	.073	[.050, .095]	.073	1.60
White men	2.36	.993	.991	.061	[.048, .074]	.042	0.68
Three-factor (Peterson et al., 2007)	13.1	.903	.890	.133	[.128, .138]	.088	3.00
Three-factor (Hilbert et al., 2012)	13.42	.903	.890	.135	[.130, .139]	.084	3.42
Female athlete three-factor (Darcy et al., 2013)	15.22	.912	.898	.147	[.142, .153]	.085	2.79
Male athlete three-factor (Darcy et al., 2013)	17.47	.891	.876	.155	[.150, .160]	.100	3.79
Three-factor (White et al., 2014)	15.82	.896	.896	.147	[.142, .152]	.093	3.66
Modified three-factor (Grilo et al., 2013)	5.20	.996	.993	.078	[.059, .099]	.030	0.18
Two-factor (Penelo et al., 2012)	15.10	.883	.870	.144	[.139, .148]	.104	4.68
Full one-factor (Pennings & Wojciechowski, 2004)	16.64	.857	.843	.151	[.147, .155]	.112	5.90
Brief one-factor (Allen et al., 2011)	37.71	.956	.938	.232	[.218, .246]	.069	0.80

Note: Fit statistics for CFA models of the Eating Disorder Examination-Questionnaire. Lavaan provides 90% CIs for RMSEA, as RMSEA tends to over-reject acceptable fitting models in smaller samples (Herzog & Boomsma, 2009). The chosen model is bolded.

square difference test was significant,  $\Delta\chi^2(8) = 25.79, p < .01$ , indicating differences in latent means across groups. Therefore, we conducted one-way ANCOVAs with post hoc Tukey tests, adjusting for BMI, to identify the nature of these observed differences.

### 3.2.1 | Post hoc analyses

Significant differences across racial groups emerged for all latent factors and the global score when adjusting for BMI: Overvaluation of Shape/Weight,  $F(2, 675) = 5.47, p < .01, \eta^2 = .02$  [.00, .03]; Appearance Concern,  $F(2, 675) = 13.62, p < .01, \eta^2 = .04$  [.02, .06]; Restraint,  $F(2, 675) = 4.50, p = .01, \eta^2 = .01$  [.00, .03]; Eating Concern,  $F(2, 675) = 8.22, p < .01, \eta^2 = .02$  [.01, .04]; global score,  $F(2, 675) = 11.11, p < .01, \eta^2 = .03$  [.01, .05]. Current BMI was also significantly associated with all latent factors ( $ps < .01$ ). Scores ranged from 0 to 6.

#### Subscale differences

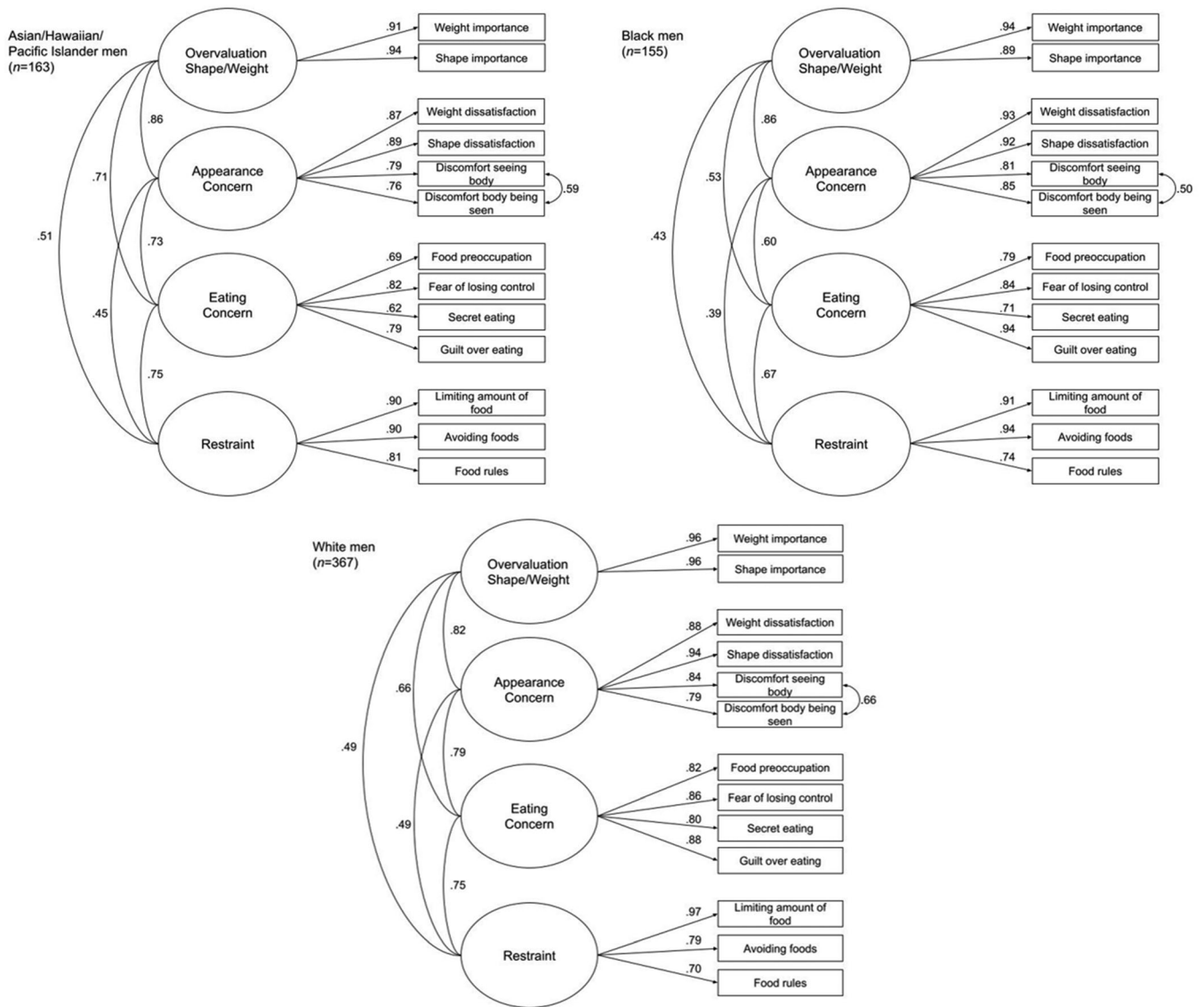
All post hoc Tukey contrasts adjusted for BMI. Black and White men did not differ significantly on Overvaluation of Shape/Weight (mean difference =  $-.09, p = .85$ ), Appearance Concern (mean difference =  $-.34, p = .07$ ), Eating Concern (mean difference =  $-.114, p = .35$ ), or the global score (mean difference =  $-.24, p = .08$ ). However, Black men reported significantly less Restraint than White men (mean difference =  $-.59, p < .01, d = .24$ ).

Asian/Hawaiian/Pacific Islander men reported significantly greater Overvaluation, Appearance Concern, Eating Concern, and

global scores than both Black (Overvaluation: mean difference =  $.60, p = .01, d = .29$ ; Appearance Concern: mean difference =  $.95, p < .01, d = .52$ ; Eating Concern: mean difference =  $.44, p < .001, d = .42$ ; global score: mean difference =  $.61, p < .001, d = .49$ ) and White men (Overvaluation: mean difference =  $.50, p = .01, d = .28$ ; Appearance Concern: mean difference =  $.61, p < .01, d = .38$ ; Eating Concern: mean difference =  $.30, p < .01, d = .28$ ; global score: mean difference =  $.37, p < .01, d = .29$ ). When adjusting for BMI, Restraint did not differ significantly between Asian/Hawaiian/Pacific Islander and White men (mean difference =  $.01, p = .99$ ), but Asian/Hawaiian/Pacific Islander men reported significantly greater Restraint than Black men (mean difference =  $.60, p = .01, d = .24$ ).

#### Differences in disordered eating behaviors

There were significant differences in the likelihood of endorsing any OBEs, LOC eating episodes, and fasting behaviors across groups. Specifically, Asian/Hawaiian/Pacific Islander and White men were significantly more likely to endorse any OBEs,  $\chi^2(2) = 11.08, p < .01$  than Black men (NNT: Asian/Hawaiian/Pacific Islander-Black men = 6.5; White-Black men = 8.8). Similarly, both groups were more likely to endorse any LOC eating episodes than Black men,  $\chi^2(2) = 9.40, p = .01$ , (NNT: Asian/Hawaiian/Pacific Islander-Black men = 8.4; White-Black men = 16.1). Groups did not differ significantly on their likelihood of endorsing purging behaviors or compensatory exercise. However, Asian/Hawaiian/Pacific Islander men were significantly more likely to endorse any fasting than White or Black men,  $\chi^2(2) = 17.24, p < .01$ , NNT = 6.0 (both comparisons).



**FIGURE 1** Factor loadings of best-fitting model for Asian, Black, and White men

**TABLE 3** Summary of EDE-Q measurement invariance of the respecified Parker et al. (2016) four-factor model

Model	$\chi^2$	df	CFI	TLI	RMSEA [90% CI]	SRMR	Model comparison	$\Delta\chi^2$	$\Delta df$	$\Delta CFI$	$\Delta TLI$	$\Delta RMSEA$	$\Delta SRMR$
M1: Configural invariance	321.92	174	.992	.990	.061 [0.051, .072]	.052	-	-	-	-	-	-	-
M2: Threshold invariance	434.03	262	.991	.992	.054 [0.045, .063]	.052	M1	100.71	88	-.001	.002	-.007	.000
M3: Loading invariance	430.25	280	.992	.994	.049 [0.039, .057]	.052	M2	12.53	18	.001	.002	-.005	.000
M4: Intercept invariance	430.14	298	.993	.995	.044 [0.035, .053]	.053	M3	2.53	18	.001	.001	-.005	.001
M5: Mean invariance	624.07	306	.984	.988	.068 [0.060, .075]	.053	M4	25.79*	8	-.009	-.007	.024	.000

Note: N = 685; Asian/Hawaiian/Pacific Islander = 163; Black = 155; White = 367. Chi-square statistics are robust estimations;  $\Delta\chi^2$  and  $\Delta df$  are adjusted differences in scaled chi-square. Configural, threshold, loading, intercept, and mean invariance refer to equivalence of the factor structure, item thresholds, item loadings, item intercepts, and latent means across groups. \*p < .01.

**TABLE 4** Parker et al.'s (2016) four-factor Model (respecified) Eating Disorder Examination-Questionnaire means, standard deviations, and endorsement of disordered eating behaviors in male university students

	Asian/Hawaiian/Pacific Islander men, <i>n</i> = 163		Black men, <i>n</i> = 155		White men, <i>n</i> = 367	
	<i>M</i> ( <i>SD</i> )	$\omega/\rho$	<i>M</i> ( <i>SD</i> )	$\omega/\rho$	<i>M</i> ( <i>SD</i> )	$\omega/\rho$
Overvaluation	2.42 (1.86) <sup>a</sup>	.81 [.75, .86]	1.88 (1.80) <sup>b</sup>	.80 [.73, .85]	1.92 (1.73) <sup>b</sup>	.86 [.83, .88]
Appearance concern	2.35 (1.70) <sup>a</sup>	.89 [.85, .92]	1.49 (1.59) <sup>b</sup>	.90 [.86, .93]	1.73 (1.61) <sup>b</sup>	.91 [.88, .93]
Eating concern	0.95 (1.11) <sup>a</sup>	.71 [.60, .80]	0.51 (0.99) <sup>b</sup>	.82 [.69, .90]	0.65 (1.06) <sup>b</sup>	.83 [.76, .88]
Restraint	1.86 (1.99) <sup>a</sup>	.85 [.78, .89]	1.40 (1.81) <sup>b</sup>	.84 [.76, .89]	1.84 (1.90) <sup>ab</sup>	.79 [.74, .83]
Global score	1.82 (1.28) <sup>a</sup>	.89 [.85, .91]	1.23 (1.14) <sup>b</sup>	.88 [.80, .92]	1.45 (1.22) <sup>b</sup>	.90 [.87, .92]
	Any occurrence	Regular occurrence	Any occurrence	Regular occurrence	Any occurrence	Regular occurrence
OBEs	30.1% <sup>a</sup>	10.4% <sup>a</sup>	14.8% <sup>b</sup>	.6% <sup>b</sup>	26.2% <sup>a</sup>	4.9% <sup>c</sup>
LOC eating	19.6% <sup>a</sup>	10.4% <sup>a</sup>	7.7% <sup>b</sup>	1.9% <sup>b</sup>	13.9% <sup>a</sup>	5.7% <sup>a,b</sup>
Self-induced vomiting	3.1%	0.7%	1.3%	0.8%	1.9%	1.4%
Laxative misuse	1.8%	–	1.3%	–	0.8%	–
Diuretic misuse	1.8%	–	1.3%	–	0.3%	–
Purging behaviors (all)	4.3% <sup>a</sup>	2.5% <sup>a</sup>	3.2% <sup>a</sup>	1.3% <sup>a</sup>	2.2% <sup>a</sup>	1.4% <sup>a</sup>
Compensatory exercise	39.9% <sup>a</sup>	15.8% <sup>a</sup>	37.4% <sup>a</sup>	10.3% <sup>a</sup>	45.0% <sup>a</sup>	12.9% <sup>a</sup>
Fasting	40.5% <sup>a</sup>	14.1% <sup>a</sup>	23.9% <sup>b</sup>	5.8% <sup>b</sup>	23.7% <sup>b</sup>	7.3% <sup>b</sup>

Note: Spearman's rank correlation coefficient is presented as the reliability estimate for the two-item overvaluation factor; McDonald's omega is presented as the reliability estimate for all other factors. Any occurrence includes proportion of men endorsing the behavior at least once in the previous 28 days; regular occurrence includes proportion of men endorsing the behavior  $\geq 4$  times in the last 28 days (Berg et al., 2012), except for compensatory exercise and fasting (going eight or more waking hours without food). Regular occurrence of compensatory exercise was  $\geq 20$  episodes and fasting  $\geq 13$  out of 28 days (Lavender et al., 2010). Purging behaviors (all) included episodes of self-induced vomiting, laxative, and diuretic misuse. Means and proportions with the same superscript letter are not significantly different (e.g., the 5.8% of Black [superscript a] and 7.3% of White men [superscript a] endorsing regular fasting did not differ significantly). Proportions with a different superscript letter indicate significant differences (e.g., the 10.4% of Asian/Hawaiian/Pacific Islander men endorsing regular LOC eating [superscript a] was significantly more than the 1.9% of endorsing Black men [subscript b]). Abbreviations: LOC eating, loss-of-control eating episodes; OBEs, objective binge episodes.

Asian/Hawaiian/Pacific Islander men were more likely than Black or White men, and White men were more likely than Black men to endorse regular OBEs,  $\chi^2(2) = 14.98, p = .01$  (NNT: Asian/Hawaiian/Pacific Islander-Black men = 10.2; Asian/Hawaiian/Pacific Islander-White men = 18.2; White-Black men = 23.3). Asian/Hawaiian/Pacific Islander men were more likely than Black men to endorse regular LOC eating,  $\chi^2(2) = 10.28, p = .01$ , NNT = 11.8. Groups did not differ in their likelihood of endorsing regular purging,  $\chi^2(2) = 0.98, p = .61$ , or regular compensatory exercise,  $\chi^2(2) = 1.86, p = .40$ . Finally, Asian/Hawaiian/Pacific Islander men were more likely than Black and White men to endorse regular fasting,  $\chi^2(2) = 8.08, p = .02$  (NNT: Asian/Hawaiian/Pacific Islander-Black = 12.0; Asian/Hawaiian/Pacific Islander-White = 14.7).

Hawaiian/Pacific Islander men scored above the clinical threshold ( $\geq 1.68$ ), established with other male samples (Schaefer et al., 2018), on the global score. Across groups, scores on the Eating Concern subscale were the lowest, whereas Overvaluation scores were the highest within each group. The most common disordered eating behavior in all groups was compensatory exercise, with any occurrence endorsed by over one-third of Black and Asian/Hawaiian/Pacific Islander men, and almost half of White men. Over 40% of Asian/Hawaiian/Pacific Islander men and almost one-quarter of Black and White men (each) endorsed any fasting. OBEs were also relatively common, occurring at least once in over one-quarter of White and Asian/Hawaiian/Pacific Islander men (each) and almost 15% of Black men. Purging behaviors were infrequently endorsed across groups (any occurrence <5%).

### 3.3 | Subscale scores and behavioral frequencies

See Table 4 for means, standard deviations, and internal consistency (McDonald's omega) of subscale scores, and behavioral item endorsement by racial/ethnic group. Internal consistency coefficients were above acceptable levels (Nájera Catalán, 2018) for the overall sample:  $\rho = .83$  (Overvaluation);  $\omega = .91$  (Appearance Concern);  $\omega = .83$  (Eating Concern);  $\omega = .81$  (Restraint);  $\omega = .88$  (global score). Asian/

## 4 | DISCUSSION

The EDE-Q is one of the most widely used research and clinical tools (Towne et al., 2017), but its four-factor structure is rarely supported across different samples (e.g., bariatric surgery candidates, clinical samples; Grilo et al., 2013; Peterson et al., 2007). Moreover, there is no known evaluation of the EDE-Q's factor structure or invariance in a sample of racially and ethnically diverse university men. To address



this gap, this study evaluated the factor structure and invariance of the EDE-Q in a large sample of Asian/Hawaiian/Pacific Islander, Black, and White American college men.

The modified three-factor structure (Grilo et al., 2013, 2015) initially provided the best fit to our data; however, given concerns about the reliability of subscales with only two items (Hair et al., 2010), as well as the substantial item reduction and potential for reduced content validity (Rand-Giovannetti et al., 2020), we evaluated MIs for the next best-fitting model. After respecifying the model to allow the error terms for two similarly worded items to correlate, and removing one item that contributed to poor fit, a respecified version of the four-factor model proposed by Parker et al. (2016) (i.e., Overvaluation of Shape/Weight, Appearance Concern, Eating Concern, and Dietary Restraint) yielded the best fit to the data. Although one factor comprised only two items, we chose to proceed with this model over the Grilo et al.'s (2013) model in favor of reliability and content validity (Hair et al., 2010). A series of multigroup CFA models provided evidence of configural, threshold, loading, and intercept invariance, suggesting any observed variability in scores was attributable to true group differences, rather than potential measurement error (Chen, 2008). Our study adds to the growing literature failing to find support for the original four-factor structure across multiple samples (e.g., Barnes et al., 2012; Peterson et al., 2007; Rand-Giovannetti et al., 2020).

More specifically, it is important to emphasize the challenges we encountered in identifying a reliable, valid, and well-fitting EDE-Q model in racially diverse college men. Of the 12 models tested, only two displayed acceptable fit, and both of these included factors comprising only two items. Given these results, researchers interested in administering the EDE-Q to racially diverse college men might consider using global and subscale scores derived from the respecified Parker et al. (2016) model, given its invariance across groups. However, given that the Overvaluation factor in this model has only two items, it should be interpreted with caution, as scales with fewer items tend to demonstrate lower internal consistency (Hair et al., 2010).

Our results also raise the question of whether the EDE-Q is the optimal measure to assess ED symptomatology in racially diverse young men, particularly given the considerable lack of research support for the original four-factor model with various groups (e.g., Peterson et al., 2007). Our findings potentially point to the need to: (a) culturally-adapt this measure for specific populations; (b) consider developing new measures that capture gender- and culturally salient constructs relevant to the population of interest (e.g., the Muscle Dysmorphic Disorder Inventory, Hildebrandt et al., 2004); and/or (c) consider developing new, more comprehensive measures that include other constructs implicated in eating pathology (e.g., negative attitudes toward obesity, Eating Pathology Symptoms Inventory; Forbush et al., 2013). Although providers might still be able to collect meaningful clinical data while using the traditional EDE-Q four-factor structure with racially diverse male clients, any interpretations of global and subscale scores should be made with caution, and bolstered by triangulation with other clinical tools (e.g., interview, medical chart review, additional eating and body image questionnaires).

When measuring ED symptoms with the respecified Parker et al. (2016) model, Asian/Hawaiian/Pacific Islander men reported significantly higher levels of ED psychopathology relative to their Black and White peers. Specifically, they reported greater Overvaluation of Shape/Weight, Appearance Concern, and Eating Concern, and were more likely to endorse OBEs and fasting episodes than their peers, although effects were generally small. These findings are consistent with prior research that reported that Asian/Hawaiian/Pacific Islander American men and boys endorsed greater ED symptomatology relative to other racial/ethnic groups of males (N. R. Kelly et al., 2015; Rodgers et al., 2017, 2018). The mechanisms underlying this observed group difference cannot be determined from either the current study, or the previous investigations cited, and should be investigated in future research.

Most ED research with Asian/Hawaiian/Pacific Islander populations has focused on women (e.g., Yu et al., 2019), and has suggested that multiple factors might contribute to ED onset and maintenance in this group, including acculturative stress, conflicting appearance ideals (e.g., Asian/Hawaiian/Pacific Islander versus Eurocentric beauty standards), cultural value conflicts (e.g., collectivistic versus individualistic), and differing attitudes and beliefs surrounding food (Goel et al., 2021; Goel, Thomas, et al., 2022; Rodgers et al., 2018; Yu et al., 2019). Current results suggest that Asian/Hawaiian/Pacific Islander college men might also be at elevated risk for ED psychopathology. However, as most observed differences were small in magnitude, further research is needed to confirm these findings. If supported, investigation into the potential mechanisms of disordered eating symptoms (and elevations) in Asian/Hawaiian/Pacific Islander college men is needed.

Our findings add to the growing literature that the profiles and patterns of ED behaviors often differ between men and women. Because symptoms in men often appear driven by internalization of a muscular ideal, men with EDs might present with higher levels of compulsive exercise and other muscularity-driven behaviors, and lower levels of purging (Murray et al., 2017). Thus, men presenting for treatment might initially be missed by providers trained in detecting “typical” ED manifestations (e.g., intense fear of weight gain, low weight). Further research is necessary given the notable increase of disordered eating in boys and men in recent years (Mitchison et al., 2014; Nagata, Ganson, et al., 2020).

The endorsement of ED behaviors in this sample of racially and ethnically diverse university men was striking. In our sample, between 37% and 45% of men reported any occurrence of compensatory exercise, which is higher than the 31% found in a sample of college men published in 2010 (Lavender et al., 2010). Endorsement of OBEs and purging behaviors were comparable to those reported in the same study, but a considerably higher proportion of Asian/Hawaiian/Pacific Islander men endorsed any occurrence of fasting (41%) than in Lavender et al. (2010) (24%). Given that this study was described to participants as addressing eating and body image, it is possible these rates are at least partially attributable to selection bias. Thus, replication in a larger, epidemiological survey would be useful to evaluate whether behaviors such as compensatory exercise and fasting are rising among college men, both across and within racial and ethnic groups.

This study has several strengths. In particular, it is one of the most comprehensive studies of the EDE-Q factor structure in men. It also includes a racially and ethnically diverse sample of college men and uses advanced statistical procedures to identify the most relevant EDE-Q factor structure for these groups. However, this study should be interpreted within the context of its limitations. First, we ultimately retained a model that had one factor with two indicators, which could limit reliability and construct validity (Hair et al., 2010). If researchers choose to proceed with the respecified Parker et al. (2016) four-factor model to assess ED symptoms in these groups in the future, we recommend collecting data to evaluate convergent, discriminant, and concurrent validity, and test-retest reliability to provide further support for this model's psychometric properties. Another limitation is that Latinx, indigenous groups, and other populations were not included (Mikhail & Klump, 2020). Furthermore, considering the heterogeneity within and across Asian/Hawaiian/Pacific Islander subgroups (Cummins et al., 2005; Goel et al., 2021; Goel, Thomas, et al., 2022) future research should assess ED behaviors within distinct Asian/Hawaiian/Pacific Islander groups. We restricted analyses to individuals self-identifying as male, limiting generalizability to other gender groups (Nagata, Ganson, et al., 2020), and our samples of Asian/Hawaiian/Pacific Islander and Black men were below conventional recommendations for CFA (Kline, 2011). Although WLSMV estimation is more powerful than other estimation methods with small sample sizes (Li, 2016), evaluation of the EDE-Q's factor structure should be replicated in larger male samples. Lastly, self-report of disordered eating behaviors may be influenced by sociocultural messaging about proper serving sizes or about the acceptability of losing control over one's behavior, and also rely on subjective self-perception. Future research can investigate the extent to which cultural socialization impacts self-report of disordered eating behaviors.

## 5 | CONCLUSIONS

Men and people of color are underrepresented in the ED literature and disparities in the prevention, assessment, and treatment of these conditions in these groups persist (Egbert et al., 2022; Goel, Thomas, et al., 2022, Murray et al., 2017). To optimize ED screening and detection, it is imperative that our most widely used measurement tools are validated within and across groups, and invariance is established to enhance confidence in the validity of score interpretations and comparisons. Current data failed to support the EDE-Q's original four-factor structure. An alternate respecified four-factor structure (Parker et al., 2016) best fit our data. Moreover, Asian/Hawaiian/Pacific Islander men reported higher levels of global ED symptomatology, relative to White and Black men. Further, endorsement of compensatory exercise was high, and fasting and OBEs were also common across all racial groups in this sample. Purging behaviors were low (<5%), consistent with past research (Murray et al., 2017). Due to the pervasive stigma surrounding EDs, especially for men and people of color, they may be less likely to seek treatment, despite significant medical complications (Griffiths

et al., 2014, 2015; Marques et al., 2011). Thus, providers working in diverse healthcare settings should prioritize assessing EDs in men using a data-informed approach to assessment and interpretation.

## ACKNOWLEDGMENTS

Neha J. Goel was supported by the National Institutes of Health (grant F31 MD015679). Carolyn Blair Burnette was supported by Award Number T32HL150452 from the National Heart, Lung, and Blood Institute (PI: Dianne Neumark-Sztainer). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the National Institute on Minority Health and Health Disparities, or the National Heart, Lung, and Blood Institute. The authors would like to thank Alexandria Davies, MS, for assistance with data collection, and all of the participants from this study.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Neha J. Goel  <https://orcid.org/0000-0002-8191-3359>

Carolyn Blair Burnette  <https://orcid.org/0000-0002-5003-5856>

## REFERENCES

- Allen, K. L., Byrne, S. M., Lampard, A., Watson, H. J., & Fursland, A. (2011). Confirmatory factor analysis of the Eating Disorder Examination-Questionnaire (EDE-Q). *Eating Behaviors*, 12(2), 143–151. <https://doi.org/10.1016/j.eatbeh.2011.01.005>
- Bardone-Cone, A. M., & Boyd, C. A. (2007). Psychometric properties of eating disorder instruments in Black and White young women: Internal consistency, temporal stability, and validity. *Psychological Assessment*, 19(3), 356–362. <https://doi.org/10.1037/1040-3590.19.3.356>
- Barnes, J., Prescott, T., & Muncer, S. (2012). Confirmatory factor analysis for the Eating Disorder Examination Questionnaire: Evidence supporting a three-factor model. *Eating Behaviors*, 13(4), 379–381. <https://doi.org/10.1016/j.eatbeh.2012.05.001>
- Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the Eating Disorder Examination and Eating Disorder Examination-Questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, 45(3), 428–438. <https://doi.org/10.1002/eat.20931>
- Boutté, R. L. (2020). Black women and body image: Implications of the racialized beauty aesthetic. (Unpublished doctoral dissertation). Virginia Commonwealth University.
- Brown, T. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
- Burnette, C. B., Boutté, R. L., Sosnowski, D. W., & Mazzeo, S. E. (2020). Measurement invariance of the SATAQ-4 in Black and White undergraduate women. *Body Image*, 32, 85–93. <https://doi.org/10.1016/j.bodyim.2019.11.004>
- Burnette, C. B., Davies, A. E., Boutté, R. L., & Mazzeo, S. E. (2019). What are you losing it for? Weight suppression motivations in undergraduates. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 25(2), 1–12. <https://doi.org/10.1007/s40519-018-00635-y>
- Burnette, C. B., & Mazzeo, S. E. (2020). Examining the contribution of weight-bias internalization to the associations between weight suppression and disordered eating in undergraduates. *Eating Behaviors*, 37, 101392. <https://doi.org/10.1016/J.EATBEH.2020.101392>

- Byrne, B. M., Oakland, T., Leong, F. T. L., van de Vijver, F. J. R., Hambleton, R. K., Cheung, F. M., & Bartram, D. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology, 3*(2), 94–105. <https://doi.org/10.1037/a0014516>
- Byrne, S. M., Allen, K. L., Lampard, A. M., Dove, E. R., & Fursland, A. (2010). The factor structure of the eating disorder examination in clinical and community samples. *International Journal of Eating Disorders, 43*(3), 260–265. <https://doi.org/10.1002/eat.20681>
- Calugi, S., Milanese, C., Sartirana, M., El Ghoch, M., Sartori, F., Geccherle, E., ... Dalle Grave, R. (2017). The Eating Disorder Examination Questionnaire: Reliability and validity of the Italian version. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity, 22*(3), 509–514. <https://doi.org/10.1007/s40519-016-0276-6>
- Carey, M., Kupeli, N., Knight, R., Troop, N. A., Jenkinson, P. M., & Preston, C. (2019). Eating Disorder Examination Questionnaire (EDE-Q): Norms and psychometric properties in U.K. females and males. *Psychological Assessment, 31*(7), 839–850. <https://doi.org/10.1037/pas0000703>
- Chan, C. W., & Leung, S. F. (2015). Validation of the eating disorder examination questionnaire: An online version. *Journal of Human Nutrition and Dietetics, 28*(6), 659–665. <https://doi.org/10.1111/jhn.12275>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*(5), 1005–1018. <https://doi.org/10.1037/a0013193>
- Cheng, Z. H., Perko, V. L., Fuller-Marashi, L., Gau, J. M., & Stice, E. (2019). Ethnic differences in eating disorder prevalence, risk factors, and predictive effects of risk factors among young women. *Eating Behaviors, 32*, 23–30. <https://doi.org/10.1016/j.eatbeh.2018.11.004>
- Compte, E. J., Nagata, J. M., Sepulveda, A. R., Schweiger, S., Sbdar, L. S., Lic Silva, B. C. S., ... Muray, S. B. (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*, 740–745. <https://doi.org/10.1002/eat.23075>
- Cummins, L. H., Simmons, A. M., & Zane, N. W. (2005). Eating disorders in Asian populations: A critique of current approaches to the study of culture, ethnicity, and eating disorders. *American Journal of Orthopsychiatry, 75*(4), 553–574. <https://doi.org/10.1037/0002-9432.75.4.553>
- Darcy, A. M., Hardy, K. K., Crosby, R. D., Lock, J., & Peebles, R. (2013). Factor structure of the Eating Disorder Examination Questionnaire (EDE-Q) in male and female college athletes. *Body Image, 10*(3), 399–405. <https://doi.org/10.1016/j.bodyim.2013.01.008>
- Dunn, T. J., Baguley, T., & Brunson, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399–412. <https://doi.org/10.1111/bjop.12046>
- Egbert, A., Hunt, R. A., Williams, K., Burke, N. L., & Jennings Mathis, K. (2022). Reporting racial and ethnic diversity in eating disorder research over the past 20 years. *International Journal of Eating Disorders, 1*–8. <https://doi.org/10.1002/eat.23666>
- Elder, K. A., & Grilo, C. M. (2007). The Spanish language version of the Eating Disorder Examination Questionnaire: Comparison with the Spanish language version of the Eating Disorder Examination and test-retest reliability. *Behaviour Research and Therapy, 45*, 1369–1377. <https://doi.org/10.1016/j.brat.2006.08.012>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370.
- Fairburn, C. G., Cooper, Z., & O'Connor, M. (1993). The eating disorder examination. *International Journal of Eating Disorders, 6*, 1–8.
- Forbush, K. T., Hilderbrand, L. A., Bohrer, B. K., & Chapa, D. A. (2019). Test-retest reliability of common measures of eating disorder symptoms in men versus women. *Assessment, 26*(3), 419–431. <https://doi.org/10.1177/1073191117700267>
- Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., ... Watson, D. (2013). Development and validation of the Eating Pathology Symptoms Inventory (EPSI). *Psychological Assessment, 25*(3), 859–878. <https://doi.org/10.1037/a0032639>
- Friborg, O., Reas, D. L., Rosenvinge, J. H., & Rø, Ø. (2013). Core pathology of eating disorders as measured by The Eating Disorder Examination Questionnaire (EDE-Q): The predictive role of a nested general and primary factors. *International Journal of Methods in Psychiatric Research, 22*(3), 195–203. <https://doi.org/10.1002/mpr.1389>
- Goel, N. J., Jennings Mathis, K., Egbert, A. H., Petteyway, F., Breithaupt, L., Eddy, K. T., ... Graham, A. K. (2022). Accountability in promoting representation of historically marginalized racial and ethnic populations in the eating disorders field: A call to action. *International Journal of Eating Disorders, XX, XX–XX*. <https://doi.org/10.1002/eat.23682>
- Goel, N. J., Thomas, B., Boutté, R. L., Kaur, B., & Mazzeo, S. E. (2021). Body image and eating disorders among South Asian American women: What are we missing? *Qualitative Health Research, 31*(13), 2512–2527. <https://doi.org/10.1177/10497323211036896>
- Goel, N. J., Thomas, B., Boutté, R. L., Kaur, B., & Mazzeo, S. E. (2022). “What will people say?”: Mental health stigmatization as a barrier to eating disorders treatment-seeking for south Asian American women. *Asian American Journal of Psychology, XX, XX–XX*. <https://doi.org/10.1037/aap0000271>
- Griffiths, S., Mond, J. M., Li, Z., Gunatilake, S., Murray, S. B., Sheffield, J., & Touyz, S. (2015). Self-stigma of seeking treatment and being male predict an increased likelihood of having an undiagnosed eating disorder. *International Journal of Eating Disorders, 48*, 775–778. <https://doi.org/10.1002/eat.22413>
- Griffiths, S., Mond, J. M., Murray, S. B., & Touyz, S. (2014). The prevalence and adverse associations of stigmatization in people with eating disorders. *International Journal of Eating Disorders, 48*, 767–774. doi: 10.1002/eat.22353
- Grilo, C. M., Henderson, K. E., Bell, R. L., & Crosby, R. D. (2013). Eating Disorder Examination-Questionnaire factor structure and construct validity in bariatric surgery candidates. *Obesity Surgery, 23*(5), 657–662. <https://doi.org/10.1007/s11695-012-0840-8>
- Grilo, C. M., Reas, D. L., Hopwood, C. J., & Crosby, R. D. (2015). Factor structure and construct validity of The Eating Disorder Examination-Questionnaire in college students: Further support for a modified brief version. *International Journal of Eating Disorders, 48*(3), 284–289. <https://doi.org/10.1002/eat.22358>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River: Prentice Hall.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Hart, L. M., Granillo, M. T., Jorm, A. F., & Paxton, S. J. (2011). Unmet need for treatment in the eating disorders: A systematic review of eating disorder specific treatment seeking among community cases. *Clinical Psychology Review, 31*(5), 727–735. <https://doi.org/10.1016/j.cpr.2011.03.004>
- Herzog, W., & Boomsma, A. (2009). Small-sample robust estimators of noncentrality-based and incremental model fit. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(1), 1–27. <https://doi.org/10.1080/10705510802561279>
- Hilbert, A., De Zwaan, M., & Braehler, E. (2012). How frequent are eating disturbances in the population? Norms of the eating disorder examination-questionnaire. *PLoS one, 7*(1), e29125. <https://doi.org/10.1371/journal.pone.0029125>

- Hilbert, A., Tuschen-Caffier, B., Karwautz, A., Niederhofer, H., & Munsch, S. (2007). Eating Disorder Examination-Questionnaire. *Diagnostica*, 53(3), 144–154. <https://doi.org/10.1026/0012-1924.53.3.144>
- Hildebrandt, T., Langenbucher, J., & Schlundt, D. G. (2004). Muscularity concerns among men: Development of attitudinal and perceptual measures. *Body Image*, 1(2), 169–181.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037//1082-989x.3.4.424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kelly, K. R., Mitchell, K. S., Gow, R. W., Trace, S. E., Lydecker, J. A., Bair, C. E., & Mazzeo, S. E. (2012). An evaluation of the reliability and construct validity of eating disorder measures in White and Black women. *Psychological Assessment*, 24(3), 608–617. <https://doi.org/10.1037/a0026457>
- Kelly, N. R., Cotter, E. W., & Mazzeo, S. E. (2012). Eating Disorder Examination Questionnaire (EDE-Q): Norms for Black women. *Eating Behaviors*, 13(4), 429–432. <https://doi.org/10.1016/j.eatbeh.2012.09.001>
- Kelly, N. R., Cotter, E. W., Tanofsky-Kraff, M., & Mazzeo, S. E. (2015). Racial variations in binge eating, body image concerns, and compulsive exercise among men. *Psychology of Men & Masculinity*, 16(3), 326–336. <https://doi.org/10.1037/a0037585>
- Kline, R. B. (2011). In R. B. Kline (Ed.), *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Kozioł, N. A., & Bovaird, J. A. (2018). The impact of model parameterization and estimation methods on tests of measurement invariance with ordered polytomous data. *Educational and Psychological Measurement*, 78(2), 272–296. <https://doi.org/10.1177/0013164416683754>
- Kraemer, H. C., Neri, E., & Spiegel, D. (2020). Wrangling with p-values versus effect sizes to improve medical decision-making: A tutorial. *International Journal of Eating Disorders*, 53(2), 302–308. <https://doi.org/10.1002/eat.23216>
- Lavender, J. M., De Young, K. P., & Anderson, D. A. (2010). Eating Disorder Examination Questionnaire (EDE-Q): Norms for undergraduate men. *Eating Behaviors*, 11(2), 119–121. <https://doi.org/10.1016/j.eatbeh.2009.09.005>
- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Luce, K. H., Crowther, J. H., & Pole, M. (2008). Eating Disorder Examination Questionnaire (EDE-Q): Norms for undergraduate women. *International Journal of Eating Disorders*, 41(3), 273–276. <https://doi.org/10.1002/eat.20504>
- Marques, L., Alegria, M., Becker, A. E., Chen, C., Fang, A., Chosak, A., & Diniz, J. B. (2011). Comparative prevalence, correlates of impairment, and service utilization for eating disorders among US ethnic groups: Implications for reducing ethnic disparities in health care access for eating disorders. *International Journal of Eating Disorders*, 44, 412–420. <https://doi.org/10.1002/eat.20787>
- Melisse, B., van Furth, E. F., & de Beurs, E. (2021). Eating Disorder Examination Questionnaire (EDE-Q): Validity and norms for Saudi nationals. *Eating and Weight Disorders*, 27, 139–150. <https://doi.org/10.1007/s40519-021-01150-3>
- Mikhail, M. E., & Klump, K. L. (2020). A virtual issue highlighting eating disorders in people of Black/African and Indigenous heritage. *International Journal of Eating Disorders*, 54, 459–467. <https://doi.org/10.1002/eat.23402>
- Mitchison, D., Hay, P., Slewa-Younan, S., & Mond, J. (2014). The changing demographic profile of eating disorder behaviors in the community. *BMC Public Health*, 14(1), 943. <https://doi.org/10.1186/1471-2458-14-943>
- Mitchison, D., & Mond, J. (2015). Epidemiology of eating disorders, eating disordered behaviour, and body image disturbance in males: A narrative review. *Journal of Eating Disorders*, 3(1), 1–9. <https://doi.org/10.1186/s40337-015-0058-y>
- Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, 42(5), 551–567. [https://doi.org/10.1016/S0005-7967\(03\)00161-X](https://doi.org/10.1016/S0005-7967(03)00161-X)
- Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., ... Mond, J. M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review*, 57, 1–11. <https://doi.org/10.1016/j.cpr.2017.08.001>
- Nagata, J. M., Compte, E. J., Murray, S. B., Schauer, R., Pak, E., Flentje, A., ... Obedin-Maliver, J. (2020). Community norms for the Eating Disorder Examination Questionnaire (EDE-Q) among cisgender bisexual plus women and men. *Eating and Weight Disorders*, 8, 1–11. <https://doi.org/10.1186/s40337-020-00352-x>
- Nagata, J. M., Ganson, K. T., & Murray, S. B. (2020). Eating disorders in adolescent boys and, eating disorders in adolescent boys young men: An update. *Current Opinion in Pediatrics*, 32(4), 476–481. <https://doi.org/10.1097/MOP.0000000000000911>
- Nagata, J. M., Murray, S. B., Compte, E. J., Pak, E. H., Schauer, R., Flentje, A., ... Obedin-Maliver, J. (2020). Community norms for the Eating Disorder Examination Questionnaire (EDE-Q) among transgender men and women. *Eating Behaviors*, 37, 101381. <https://doi.org/10.1016/j.eatbeh.2020.101381>
- Nagata, J. M., Murray, S. B., Flentje, A., Compte, E. J., Schauer, R., Pak, E., ... Obedin-Maliver, J. (2020). Eating disorder attitudes and disordered eating behaviors as measured by the Eating Disorder Examination Questionnaire (EDE-Q) among cisgender lesbian women. *Body Image*, 34, 215–220. <https://doi.org/10.1016/j.bodyim.2020.06.005>
- Nájera Catalán, H. E. (2018). Reliability, population classification and weighting in multidimensional poverty measurement: A Monte Carlo study. *Social Indicators Research*, 142(3), 887–910. <https://doi.org/10.1007/S11205-018-1950-Z>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Otani, M., Hiraide, M., Horie, T., Mitsui, T., Yoshida, T., Takamiya, S., ... Yoshiuchi, K. (2021). Psychometric properties of the Eating Disorder Examination-Questionnaire and psychopathology in Japanese patients with eating disorders. *International Journal of Eating Disorders*, 54, 203–211. <https://doi.org/10.1002/eat.23452>
- Parker, K., Mitchell, S., O'Brien, P., & Brennan, L. (2016). Psychometric evaluation of disordered eating measures in bariatric surgery candidates. *Obesity Surgery*, 26(3), 563–575. <https://doi.org/10.1007/s11695-015-1780-x>
- Penelo, E., Negrete, A., Portell, M., & Raich, R. M. (2013). Psychometric properties of the Eating Disorder Examination Questionnaire (EDE-Q) and norms for rural and urban adolescent males and females in Mexico. *PLoS One*, 8(12), e83245. <https://doi.org/10.1371/journal.pone.0083245>
- Penelo, E., Villarreal, A. M., Portell, M., & Raich, R. M. (2012). Eating Disorder Examination Questionnaire (EDE-Q): An initial trial in Spanish male undergraduates. *European Journal of Psychological Assessment*, 28(1), 76–83. <https://doi.org/10.1027/1015-5759/a000093>
- Pennings, C., & Wojciechowski, F. L. (2004). Kort Instrumenteel De Eating Disorder Examination Questionnaire (EDE-Q): Nederlandse normscores voor anorexiapatiënten en een nieteetstoornis controlegroep. - PsycNET. *Gedragstherapie*, 37(4), 293–301. <https://psycnet.apa.org/record/2005-00059-004>



- Perez, M., & Warren, C. S. (2013). Assessing eating pathology in Hispanic Americans. In L. T. Benuto (Ed.), *Guide to psychological assessment with Hispanics* (pp. 201–214). Boston, MA: Springer.
- Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., ... Le Grange, D. (2007). Psychometric properties of the Eating Disorder Examination-Questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders, 40*(4), 386–389. <https://doi.org/10.1002/eat.20373>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- R Core Team (2018). R: A language and environment for statistical computing (3.5.1). R Foundation for Statistical Computing. <http://www.r-project.org/>
- Rand-Giovannetti, D., Cicero, D. C., Mond, J. M., & Latner, J. D. (2020). Psychometric properties of the Eating Disorder Examination-Questionnaire (EDE-Q): A confirmatory factor analysis and assessment of measurement invariance by sex. *Assessment, 27*(1), 164–177. <https://doi.org/10.1177/1073191117738046>
- Rodgers, R. F., Berry, R., & Franko, D. L. (2018). Eating disorders in ethnic minorities: An update. *Current Psychiatry Reports, 20*(10), 90. <https://doi.org/10.1007/s11920-018-0938-3>
- Rodgers, R. F., Watts, A. W., Austin, S. B., Haines, J., & Neumark-Sztainer, D. (2017). Disordered eating in ethnic minority adolescents with overweight. *International Journal of Eating Disorders, 50*(6), 665–671. <https://doi.org/10.1002/eat.22652>
- Rose, J. S., Vaewsorn, A., Rosselli-Navarra, F., Wilson, G. T., & Weissman, R. S. (2013). Test-retest reliability of the Eating Disorder Examination-Questionnaire (EDE-Q) in a college sample. *Journal of Eating Disorders, 1*(1), 42. <https://doi.org/10.1186/2050-2974-1-42>
- Rosseel, Y. (2012). lavaan: A R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(2), 167–180. <https://doi.org/10.1080/10705511.2014.882658>
- Schaefer, L. M., Smith, K. E., Leonard, R., Wetterneck, C., Smith, B., Farrell, N., ... Thompson, J. K. (2018). Identifying a male clinical cutoff on the Eating Disorder Examination-Questionnaire (EDE-Q). *International Journal of Eating Disorders, 51*, 1357–1360. <https://doi.org/10.1002/eat.22972>
- Schaumberg, K., Welch, E., Breithaupt, L., Hübel, C., Baker, J. H., Munn-Chernoff, M. A., ... Bulik, C. M. (2017). The science behind the Academy For Eating Disorders' nine truths about eating disorders. *European Eating Disorders Review, 25*(6), 432–450. <https://doi.org/10.1002/erv.2553>
- Serier, K. N., Smith, J. E., & Yeater, E. A. (2018). Confirmatory factor analysis and measurement invariance of the Eating Disorder Examination Questionnaire (EDE-Q) in a non-clinical sample of non-Hispanic White and Hispanic women. *Eating Behaviors, 31*, 53–59. <https://doi.org/10.1016/j.eatbeh.2018.08.004>
- Stunkard, A. J., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research, 29*(1), 71–83. [https://doi.org/10.1016/0022-3999\(85\)90010-8](https://doi.org/10.1016/0022-3999(85)90010-8)
- Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 111–130. <https://doi.org/10.1080/10705511.2019.1602776>
- Towne, T. L., De Young, K. P., & Anderson, D. A. (2017). Trends in professionals' use of eating disorder assessment instruments. *Professional Psychology: Research and Practice, 48*(4), 243–250. <https://doi.org/10.1037/pro0000067>
- Treasure, J., Duarte, T. A., & Schmidt, U. (2020). Eating disorders. *Lancet, 395*(10227), 899–911. [https://doi.org/10.1016/S0140-6736\(20\)30059-3](https://doi.org/10.1016/S0140-6736(20)30059-3)
- Virginia Commonwealth University (2021). VCU Fact Card 2020-21. <https://irds.vcu.edu/media/decision-support/pdf/fact-cards/FinalADAweb2-10-21.pdf>
- Ward, Z. J., Rodriguez, P., Wright, D. R., Austin, S. B., & Long, M. W. (2019). Estimation of eating disorders prevalence by age and associations with mortality in a simulated nationally representative US cohort. *JAMA Network Open, 2*(10), e192925. <https://doi.org/10.1001/jamanetworkopen.2019.12925>
- White, H. J., Haycraft, E., Goodwin, H., & Meyer, C. (2014). Eating Disorder Examination Questionnaire: Factor structure for adolescent girls and boys. *International Journal of Eating Disorders, 47*(1), 99–104. <https://doi.org/10.1002/eat.22199>
- Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika, 81*(4), 1014–1045. <https://doi.org/10.1007/s11336-016-9506-0>
- Yu, K., Pope, S. C., & Perez, M. (2019). Clinical treatment and practice recommendations for disordered eating in Asian Americans. *Professional Psychology: Research and Practice, 50*, 279–287. <https://doi.org/10.1037/pro0000244>
- Zhang, Z. (2018). In Z. Zhang, & K.-H. Yuan (Eds.), *Practical statistical power analysis using webpower and R*. Granger, IN: ISDSA Press.

**How to cite this article:** Goel, N. J., Burnette, C. B., Weinstock, M., & Mazzeo, S. E. (2022). Eating Disorder Examination-Questionnaire: Evaluating factor structures and establishing measurement invariance with Asian/Hawaiian/Pacific Islander, Black, and White American college men. *International Journal of Eating Disorders, 55*(4), 481–493. <https://doi.org/10.1002/eat.23696>