Article:
Carey, Mark, Kupeli, Nuriye, Knight, Ruth et al. (3 more authors) (2019) Eating Disorder Examination Questionnaire (EDE-Q) : Norms and psychometric properties in U.K. females and males. Psychological Assessment. ISSN 1939-134X

https://doi.org/10.1037/pas0000703

Eating Disorder Examination Questionnaire (EDE-Q): Norms and Psychometric Properties in UK females and males

Mark Carey[1*], Nuriye Kupeli[2], Ruth Knight[1], Nicholas A. Troop[3], Paul M. Jenkinson[3†], and Catherine Preston[1†]

[1] Department of Psychology, University of York, York, United Kingdom

[2] Marie Curie Palliative Care Research Department, University College London, Division of Psychiatry, United Kingdom

[3] School of Life and Medical Sciences, University of Hertfordshire, Hatfield, United Kingdom

† These authors have shared senior authorship

***Corresponding author:**

Mark Carey, Department of Psychology, University of York, Heslington, York, YO10 5DD, United Kingdom.

Email: mac541@york.ac.uk

Declarations of Interest: None

Abstract

The Eating Disorder Examination Questionnaire (EDE-Q) is a widely used assessment of eating disorder psychopathology; however, EDE-Q norms are yet to be provided within a non-clinical UK adult sample. Secondly, there is considerable disagreement regarding the psychometric properties of this measure. Several alternative factor structures have been previously proposed, but very few have subsequently validated their new structure in independent samples and many are often confined to specific sub-populations. Therefore, in the current study, we provide norms of the original four-factor EDE-Q structure, and subsequently assess the psychometric properties of the EDE-Q in females and males using a large non-clinical UK sample (total *N*=2459). EDE-Q norms were consistently higher in females compared with males across all samples. Initial Confirmatory Factor Analyses (CFA) did not support the original four-factor structure for females or males (Phase 1). However, subsequent Exploratory Factor Analyses (EFA) revealed a three-factor structure as being the optimal fit for both females and males, using an 18-item and 16-item model, respectively (Phase 2). For females, the newly-proposed 18-item structure was validated within an independent student sample and further validated in an additional non-student sample. The 16-item three-factor male structure was also validated within an independent non-student sample, but was marginally below accepted fit indices within an independent student sample (Phase 3). Taken together, the above findings suggest that the EDE-Q factor structure may require further reassessment, with greater focus on the qualitative differences in interpretation of EDE-Q items between females and males.

Keywords: Eating Disorder Examination Questionnaire; Psychopathology; Sex Differences; Norms; Factor Analysis

Public Significance Statement

The present study suggests that the measures used to investigate eating disorder psychopathology in the non-clinical population may require reassessment in accordance with updated eating disorder symptomology. Additionally, given that this measure was originally developed using female populations, the utility of the EDE-Q measuring male eating disorder symptomology must be further considered.

Introduction

The Eating Disorder Examination Questionnaire (EDE-Q) (Fairburn & Beglin, 1994) is a well-established assessment of eating disorder (ED) psychopathology, and is widely used in both clinical and non-clinical populations. This self-report measure is derived from the Eating Disorder Examination (EDE) interview (Fairburn & Cooper, 1993), which is considered to be the gold standard in clinical ED assessment (Guest, 2000). The EDE-Q has traditionally been viewed as a reliable and valid alternative tool for identifying those most at risk for an ED (Berg, Peterson, Frazier, & Crow, 2012; Mond et al., 2008), with researchers and clinicians deeming it more cost-effective than the EDE, taking approximately 15 minutes to complete whilst maintaining a comparable degree of accuracy in ED psychopathology assessment (Mond, Hay, Rodgers, & Owen, 2006).

The availability of contextually relevant (i.e. specific to a given country and/or culture) normative data is vital for appropriate interpretation of assessments such as the EDE-Q (Welch, Birgegard, Parling, & Ghaderi, 2011). Indeed, EDE-Q norm data in non-clinical populations are important to gather to statistically determine the accurate use of this measure as a clinical screening instrument (Mond et al., 2006), and trace changes in the trajectory of body image attitudes over time and between cultures (Cash, Morrow, Hrabosky, & Perry, 2004; Karazsia, Murnen, & Tylka, 2017). Moreover, gathering norm data at the non-clinical, population level can help to inform health-related programmes and identify disordered eating behaviours which occur *prior* to adverse health outcomes associated with a clinical diagnosis (Mond, Mitchison, & Hay, 2013) Indeed, given the strong association that disordered eating behaviours have with physical and psychological well-being (Cash, 2004), researchers argue that the largest long-term health burden exists at the general population level (Mond, Hay, Rodgers, & Owen, 2009), further highlighting the importance of examining norms within non-clinical samples.

EDE-Q norms have been investigated in several non-clinical Western samples (e.g. Luce, Crowther, & Michele Pole, 2008; Machado et al., 2014; Mond et al., 2006; Villarroel, Penelo, Portell, & Raich, 2011), however, norm data for the adult UK population are lacking and are currently restricted to early adolescent samples only (Carter, Stewart, & Fairburn, 2001; White, Haycraft, Goodwin, & Meyer, 2014). Although body ideals and attitudes in the UK may appear similar to other Western cultures in females (Bell & Dittmar, 2011; Robinson & Aveyard, 2017) and males (Bazzini, Pepper, Swofford, & Cochran, 2015), research has indicated that international differences can exist within Western culture. For example, body image concerns were shown to differ between US and UK samples, with US individuals more likely to engage in self-accepting body talk compared with UK individuals (Payne, Martz, Tompkins, Petroff, & Farrow, 2011). This suggests that ED behaviours and attitudes are likely to show subtle differences between countries with the same Western culture (Luce et al., 2008). Indeed, evidence has shown that EDE-Q norm scores may vary across countries and cultures (Mond et al., 2006; Welch et al., 2011). Therefore, it is important that researchers and clinicians have up-to-date, normative EDE-Q data for the specific country within which it is used, to accurately interpret the scores of a certain individual or group (AERA et al., 1999). Furthermore, research must continue to evaluate the efficacy of the EDE-Q in assessing ED symptomology in females and males independently, given the fundamental differences in body image concerns and body ideals between sexes (e.g. Jennings & Phillips, 2017; Lavender, De Young, & Anderson, 2010; Mantilla & Birgegard, 2016; Mond et al., 2014; Smith et al., 2017) and changes to ED diagnoses under the latest DSM-5 criteria (American Psychiatric Association, 2013).

Secondly, despite its wide use as a measure of ED psychopathology in non-clinical and clinical samples (Aardoom, Dingemans, Slof Op't Landt, & Van Furth, 2012; Mond et al.,

2006), a limitation of the EDE-Q is the fact that it is simply a derivative of the original EDE clinician interview, with its original factor structure lacking empirical support. The original EDE-Q was proposed as a four-factor structure containing the subscales: *Restraint*, *Eating Concern*, *Shape Concern*, and *Weight Concern* (Fairburn & Beglin, 1994). However, as this factor structure was not empirically established, replication of the original factor structure is, unsurprisingly, limited (Forsén Mantilla, Birgegård, & Clinton, 2017). Moreover, the suitability of the factor structure may not be comparable between sexes, given the discrepancy in body ideals and body image concerns between females and males (Rand-Giovannetti, Cicero, Mond, & Latner, 2017; Smith et al., 2017). Indeed, as the EDE-Q was originally developed using female populations, this measure may not accurately reflect the current body ideals in males, such as increased muscularity and leanness (Jennings & Phillips, 2017), which influences ED psychopathology in a qualitatively different manner compared with females (Mitchison & Mond, 2015). Subsequently, further research is required to assess whether the current EDE-Q structure is successfully capturing ED psychopathology equally in females and males (Rand-Giovannetti et al., 2017). With increasing understanding of EDs, changes in ED diagnoses (American Psychiatric Association, 2013), and differences in the presentation of disordered eating behaviours and attitudes between females and males, it is important to continually evaluate and update the EDE-Q as an assessment measure.

The above points are exemplified by empirical research which has assessed the psychometric properties of the EDE-Q and failed to support the original four-factor structure using confirmatory factor analysis (CFA) within non-clinical and clinical samples (see Rand-Giovannetti et al., 2017, for review of factor structure studies). Indeed, several studies have used exploratory factor analysis (EFA) to propose alternative factor structures by removing items that load poorly onto any one factor (Rand-Giovannetti et al., 2017). Whilst the present study does not provide an exhaustive review of all alternative factor structures, studies which

retain all 22 original subscale items include a three-factor structure (Shape Concern and Weight Concern combined) (Peterson et al., 2007), a two-factor structure (Eating Concern, Shape Concern, and Weight Concern combined) (Becker et al., 2010; Penelo, Negrete, Portell, & Raich, 2013), and a one-factor (Global EDE-Q) structure (Pennings & Wojciechowski, 2004) within clinical ED and community samples. This highlights the equivocal reliability and apparent inconsistency of the current EDE-Q scoring system. However, very few studies which propose alternative EDE-Q structures have validated such structures within an independent sample (Friborg, Reas, Rosenvinge, & Rø, 2013; Grilo et al., 2010; Hrabosky et al., 2008; Kliem et al., 2016), which may limit the external validity of proposed factor structures. Moreover, whilst the suitability of the factor structure has been investigated in both female and male samples independently (e.g. Darcy, Hardy, Crosby, Lock, & Peebles, 2013), research which statistically compares the EDE-Q structure between sexes is scarce (Grilo, Reas, Hopwood, & Crosby, 2015; Kliem et al., 2016; Penelo et al., 2013; Rand-Giovannetti et al., 2017).

Given the current limitations as outlined above, the first aim of the present study was to provide EDE-Q norms of the original, four-factor structure across a large, non-clinical sample of UK females and males. A second aim was to assess the suitability of the original four-factor structure (Fairburn & Beglin, 1994), plus alternatively proposed three-factor (Peterson et al., 2007), two-factor (Becker et al., 2010; Penelo et al., 2013), and one-factor structure (Pennings & Wojciechowski, 2004) of the EDE-Q in females and males independently, using a CFA (Phase 1). Furthermore, we compared the above factor structures between female and male samples using measurement invariance analysis, to assess whether the structures are statistically equivalent between sexes. Based upon previous research in Western samples (e.g. Lavender et al., 2010; Penelo et al., 2013; Quick & Byrd-Bredbenner, 2013), we expected females to display higher EDE-Q norm scores compared with males.

Further, in line with previous research (e.g. Darcy et al., 2013; White et al., 2014), we hypothesized that CFA would fail to support the original four-factor structure for both sexes, with a poorer model fit amongst males compared with females within all of the above factor structures, given the qualitative difference in ED pathology in males (Mitchison & Mond, 2015). Following our hypothesised outcome for lack of support for previously proposed structures, we therefore conducted an EFA to obtain an optimal model fit of the EDE-Q data, in females and males independently (Phase 2). Newly-proposed factor structures were then submitted to a subsequent CFA using independent student and non-student samples, in order to validate and examine the external validity of the new structures within the broader UK population (Phase 3).

Method

Participants

The EDE-Q was assessed in a total of 2459 participants across three independent samples; two student samples (Samples 1 and 2) and one non-student sample (Sample 3). Student samples were recruited via internal university participation schemes, and the non-student sample was recruited via email, online social networking sites, and health and well-being forums. The study received departmental ethical approval and was conducted in accordance with the Declaration of Helsinki. Across all samples, participants whose age was ≥ 2 standard deviations (SD) above the sample mean were removed prior to analysis, to maintain homogeneity within each sample of females and males, respectively. The study was undertaken as follows:

Phase 1: EDE-Q norms and Initial Confirmatory Factor Analysis

1075 student participants (Sample 1) were recruited to provide EDE-Q norms for females and males based on the originally proposed four-factor structure. This sample contained 851 females (Mean age = 19.77, SD ± 1.73, Range = 17-29) and 224 males (Mean age = 20.34, SD ± 2.69, Range = 17-30). Table 1 summarizes demographic information for this sample (age, gender, BMI, and EDE-Q subscale/global scores).

Confirmatory factor analyses were also conducted in Sample 1, in females and males independently, to assess the adequacy of the original four-factor EDE-Q structure (Fairburn & Beglin, 1994), and alternative three-factor (Peterson et al., 2007), two-factor (Becker et al., 2010; Penelo et al., 2013), and one-factor (Pennings & Wojciechowski, 2004) structures.

Phase 2: Exploratory Factor Analysis

Following our initial CFA, we conducted exploratory factor analysis (EFA) on the same sample of student participants (Sample 1), to explore alternative factor solutions which provide a better fit for the EDE-Q data. This approach follows previous studies in non-clinical (Darcy et al., 2013; Forsén Mantilla et al., 2017) and clinical samples (Parker, Mitchell, O'Brien, & Brennan, 2015, 2016), where existing factor structures were not supported by an initial CFA. We used Sample 1 to conduct two separate EFAs which explored the EDE-Q structure independently in females (*N*=851) and males (*N*=224).

Phase 3: Confirmatory Factor Analysis of Newly-Proposed Factor Structures

To validate the newly-proposed female and male factor structures established from the EFA in Phase 2, further CFAs were conducted in two independent samples, comprising students (Sample 2, *N*=653) and non-students (Sample 3, *N*=731). Sample 2 (student sample) contained 489 females (Mean age = 22.16, SD ± 3.88, Range = 18 – 37) and 164 males (Mean age= 22.86, SD ± 3.69, Range = 18 – 33). Sample 3 (non-student sample) contained 561 females (Mean age = 32.68, SD ± 10.25, Range = 18 - 58) and 170 males (Mean age = 34.39, SD ± 11.08, Range = 18 - 61). Demographic information for females and males in Sample 2 and Sample 3 (age, gender, BMI, and EDE-Q subscale/global scores of original four-factor structure) are included in Supplementary Materials (Tables S6 and S7).

Measures

Eating Disorder Examination Questionnaire (EDE-Q)

The EDE-Q is a 28-item self-report questionnaire of ED psychopathology (Fairburn & Beglin, 1994). The questionnaire assesses disordered eating attitudes and behaviours within the past 28 days, consisting of four subscales: Restraint (5 items), Eating Concern (5 items), Shape Concern (8 items), and Weight Concern (5 items). A global score is calculated from the average

of the four subscales. Items are rated along a 7-point Likert scale, ranging from 0 to 6, in which higher scores signify higher ED psychopathology. This scoring is with the exception of six items which assess the frequency of ED behaviours within the past 28 days (see Table S3). These six items do not contribute to the above subscales, but do provide important information regarding overall, core disordered eating behaviours (e.g. self-induced vomiting, excessive exercise) (Fairburn & Beglin, 1994; Quick & Byrd-Bredbenner, 2013). Overall, the EDE-Q subscales and global measure have shown good internal consistency, with Cronbach's alpha ranging from .78 to .93 in non-clinical samples (Berg et al., 2012; Peterson et al., 2007).

Procedure

Participants across all three samples were directed to an online webpage wherein they completed the EDE-Q. The questionnaire was administered using online Qualtrics survey software and took participants approximately 15 minutes to complete. Demographic information including age and sex was acquired, plus height and weight which was used to calculate body mass index (BMI) (kg/m^2). The questionnaire was presented such that participants could not skip past individual items, ensuring there were no missing data. A validity item was also embedded in the survey (i.e. *"To ensure that you are paying attention, please choose agree for this question"* (Dakanalis, Zanetti, Riva, & Clerici, 2013)), with no incorrect responses reported.

Data Analysis

EDE-Q norms (Samples 1-3) are presented as the mean and standard deviation (SD) of all attitudinal EDE-Q subscale and global scores, for females and males independently. Independent samples t-tests were conducted to calculate differences between sexes on subscale and global scores (see Table 1 and Supplementary Materials - Tables S6 & S7). EDE-Q percentile ranks were calculated in addition to internal consistency using Cronbach's alpha coefficient (α) for females and males, respectively (see Supplementary Materials - Tables S1 & S2). Frequency of disordered eating behaviours were also calculated based on the diagnostic items that are independent from the EDE-Q subscales, with chi-square (χ^2) and Fisher's exact tests conducted to calculate differences in the proportion of reported disordered eating behaviours between females and males (see Supplementary Materials - Table S3).

Confirmatory factor analysis was conducted using AMOS software (Arbuckle, 2014; Version 23.0), to assess the goodness of fit for each factor structure in females and males independently. A model may be regarded as an acceptable fit if the Goodness of Fit Index (GFI), Normed Fit Index (NFI) and Comparative Fit Index (CFI) are all above .90; Adjusted Goodness of Fit Index (AGFI) is above .80 (Byrne, 1994), and Root Mean Square Error of Approximation (RMSEA) is below .10 (Browne & Cudeck, 1993). If the chi-square test (χ^2) is non-significant, the model can be regarded as acceptable, with lower statistics for the ratio of chi-square to degrees of freedom (χ^2/df) indicative a better model fit (Browne & Cudeck, 1993). For each of the four previously proposed models (Phase 1), measurement invariance was also calculated between sexes, to examine whether the factor structure presented as equivalent for females and males (Cheung & Rensvold, 2002; see Supplementary Materials S5).

For data that revealed a poor fit in the initial CFAs (Phase 1), subsequent exploratory factor analysis (EFA) was conducted using principal axis factoring (PAF) with oblique

(Promax) rotation (Phase 2). Examination of Kaiser's criterion (Kaiser, 1961) with eigenvalues above 1, in conjunction with Horn's Parallel Analysis (PA; Horn, 1965) provided a robust method in determining the optimal number of extracted factors for both sexes (Watkins, 2005). Items loading below .40, or cross-loading items of .32 or above (Forsén Mantilla et al., 2017; Tabachnick, Fidell, & Osterlind, 2001) were removed from further analyses.

Results

Phase 1: EDE-Q norms and Initial Confirmatory Factor Analysis

EDE-Q Norms and Descriptive Statistics (Sample 1)

Means and standard deviations for all original, four-factor subscale and global EDE-Q scores, and descriptive data are presented in Table 1. Independent samples t-tests revealed mean subscale and global scores as significantly higher (p <.001) for females compared with males. An independent samples t-test revealed females' mean age as significantly lower than males, with no significant difference between sexes for BMI (see Table 1). Percentile ranks, and clinical significance cut-offs are reported in Supplementary Materials (Table S1, S2, & S4). Cronbach's alpha coefficients were acceptable ($\alpha > .70$) across all subscales and global score for both sexes (see Supplementary Materials - Tables S1 & S2). Percentages of females and males who reported 'any' or 'regular' occurrence of disordered eating behaviours are presented in Supplementary Materials (Table S3). Chi-square or Fisher's exact tests showed that significantly more females reported self-induced vomiting and laxative misuse (any occurrence) compared with males. Moreover, significantly more females reported regular occurrence of objective binge episodes and dietary restraint compared with males (see Table S3).

INSERT TABLE 1

Initial Confirmatory Factor Analysis

The EDE-Q factor structure was assessed independently for females and males (Sample 1) using a CFA. As the assumption of multivariate normality was not met, maximum-likelihood (ML) estimation was used and the data were bootstrapped (see Table 2 for fit indices).

The original four-factor model containing the 22 attitudinal items was shown to be invalid for both samples due to Heywood cases, in which the standardized regression weights were larger than 1 for loadings onto item 8 *(Preoccupation with Shape or Weight)*. This was therefore treated as a specification error, and item 8 was removed from further analysis within the four-factor model. The four-factor model provided a poor fit to the data for both females and males (Table 2), with all fit indices below the accepted threshold, and a significant chi-square statistic (Browne & Cudeck, 1993). CFA was also undertaken for a three-factor (Shape Concern and Weight Concern combined), two-factor (Eating Concern, Shape Concern, and Weight Concern combined) and one-factor (Global EDE-Q) model. Similarly, the fit for all alternative models was unacceptable for females and males, with minimal change to fit indices (see Table 2). Such inflated chi-square values may be caused by large sample sizes as present for both sexes in the current study (Ullman, 2001). However, examination of the fit indices does indicate a poor data fit for each of the models tested.

Model measurement invariance analysis was undertaken to determine whether the EDE-Q factor structure was equivalent between female and male samples. All four models revealed significant differences, suggesting that female and male respondents may be interpreting EDE-Q items in a conceptually different manner. See Supplementary Materials (Table S5) for full analysis details.

INSERT TABLE 2

Phase 2: Exploratory Factor Analysis

As the data provided an inadequate fit for all previously proposed models using a CFA, a subsequent exploratory factor analysis (EFA) was conducted independently for both females and males (Sample 1) on all 22 attitudinal EDE-Q items.

Females

Parallel analysis confirmed that a three-factor model would be the optimal fit for female data. Items 6 *(Flat Stomach)* and 10 (*Fear of Weight Gain*) did not adequately load onto any factor (<.40), and items 12 (*Desire to Lose Weight*) and 21 *(Social Eating)* showed high cross-loadings. Thus, these four items were removed from analysis, meaning a PAF was re-run with Promax rotation on 18 items. Kaiser-Meyer-Olkin measure verified sampling adequacy for the analysis (KMO = .92), and Bartlett's Test of Sphericity was significant ($p < .001$). This model cumulatively explained 66.26% of the variance (see Table 3). Factor one was comprised of items related to Shape Concern and Weight Concern subscales, with the addition of one item (Item 2 - *Guilt about Eating*) from the Eating Concern subscale. Accordingly, this factor was termed *Shape and Weight Concern*. Factor two was comprised of items related to a preoccupation, and Eating Concern, with the addition of two items (Item 5- *Empty Stomach* and Item 2- *Avoidance of Eating*). This factor was termed *Preoccupation and Eating Concern*. Factor three was comprised of items related to dietary restriction and was termed *Restriction*.

Males

A three-factor model was also found to be the optimal fit for the male data. Iterative analyses were made from the original 22 items, with items 2 *(Avoidance of Eating)*, 5 (*Empty Stomach*), 6, *(Flat Stomach)*, 9 *(Fear of losing control over eating),* and 10 (*Fear of Weight Gain*) removed due to not adequately loading onto any factor (<.40), and item 12 (*Desire to Lose Weight*) removed due to cross-loading, meaning a PAF was re-run with Promax rotation

on 16 items. Kaiser-Meyer-Olkin measure verified sampling adequacy for the analysis (KMO=.91), and Bartlett's Test of Sphericity was significant ($p < .001$), with the model cumulatively explaining 67.38% of the variance (see Table 3). Similar to the factor structure found in the female sample, factor one was comprised of items related to Shape Concern and Weight Concern subscales, thus termed *Shape and Weight Concern*. Factor two was comprised of items related to a preoccupation, and Eating Concern, thus termed *Preoccupation and Eating Concern*. Factor three was comprised of items related to dietary restriction, thus also termed *Restriction*. Therefore, the factor structure was replicated in female and male samples, albeit that fewer items were retained in the factor solution for males.

INSERT TABLE 3

Phase 3: Confirmatory Factor Analysis of Newly-Proposed Factor Structures

Student Sample (Sample 2)

Consistent with our initial student sample (Sample 1), independent samples t-tests revealed mean subscale and global EDE-Q norm scores for the original, four-factor structure as significantly higher for females compared with males ($p < .001$) within an independent, non-clinical student sample (Sample 2) (see Supplementary Materials – Table S6). Moreover, EDE-Q scores are provided based on the newly-proposed, three-factor structure for females and males, respectively (see Supplementary Materials – Tables S8 & S9).

A subsequent CFA was run to evaluate the newly-proposed three-factor models established from the EFA using an independent, student sample (Sample 2). The assumption of multivariate normality was not met for either sample, therefore maximum-likelihood (ML) estimation was used and the data were bootstrapped. Within the new female sample, the newly-

proposed three-factor model showed improved fit indices compared with the previously assessed EDE-Q models, although fit statistics remained marginally outside of the accepted threshold (Browne & Cudeck, 1993). However, follow-up analyses were undertaken to assess parameters with modification indices above 10.00 (Heene, Hilbert, Harald Freudenthaler, & Bühner, 2012), with several highly correlated error terms within the same factor subsequently co-varied within the model, before re-running the analyses (see Supplementary materials Figure S1a for model covariances). Such modification of the model significantly improved the model fit, with necessary fit indices above the accepted .90 threshold and AGFI above .80 (Hu & Bentler, 1999) (see Table 4).

CFA analysis was also undertaken within the new male sample, using the newly-proposed three-factor model. Results showed similarly improved fit compared with previously assessed EDE-Q models, yet this also remained below the accepted threshold (see Table 4). Modification indices were similarly assessed, with necessary co-variances made between error terms (see Supplementary Materials Figure S1b), and further improvements made to the model, yet this did not reach the necessary threshold for good model fit (Browne & Cudeck, 1993).

Non-student Sample (Sample 3)

Consistent with Sample 1 and Sample 2, independent samples t-tests revealed mean subscale and global EDE-Q norm scores for the original, four-factor structure as significantly higher for females compared with males ($p < .001$) within an independent, non-clinical, non-student sample (Sample 3) (see Supplementary Materials – Table S7). EDE-Q scores are also provided based on the newly-proposed, three-factor structure for females and males, respectively (see Supplementary Materials – Tables S8 & S9).

A further CFA was run to evaluate the newly-proposed models established from the EFA using an independent, non-student sample (Sample 3) of non-clinical females and males, respectively. Once again, the assumption of multivariate normality was not met for either sample, therefore maximum-likelihood (ML) estimation was used and the data were bootstrapped. Within both non-student samples, the newly-proposed three-factor model showed similar fit statistics, which remained marginally outside of the accepted threshold (Browne & Cudeck, 1993). However, follow-up analyses were once again undertaken based on modification indices and further improvements were made to both models (see Supplementary Materials Figure S1c & S1d). Such modifications significantly improved the model fit in both female and male models, with necessary fit indices above the accepted .90 threshold and AGFI above .80 (Hu & Bentler, 1999) (see Table 4).

INSERT TABLE 4

Discussion

The present study provided EDE-Q norms of the originally proposed four-factor structure, and assessed the psychometric properties of the EDE-Q factor structure within a non-clinical UK sample of females and males. Given the inconsistency of previously reported psychometric properties of the EDE-Q, the present study conducted a CFA to test the original (Fairburn & Beglin, 1994) and alternatively proposed (Becker et al., 2010; Penelo et al., 2013; Pennings & Wojciechowski, 2004; Peterson et al., 2007) EDE-Q structures (Phase 1), followed by an EFA to determine an optimal fit for both female and male student samples (Phase 2). Importantly, the present study aimed to validate our newly-proposed female and male factor structures by conducting subsequent CFAs within independent student and non-student samples (Phase 3).

EDE-Q norm scores of the originally proposed four-factor structure were shown to be significantly higher amongst UK females compared with males, within both student and non-student samples. Such findings are consistent with previously published non-clinical norms comparing between sexes in Western samples (Penelo et al., 2013; Quick & Byrd-Bredbenner, 2013). Indeed, to the authors' knowledge, the present findings are the first to provide EDE-Q norms in a non-clinical UK adult sample, with global scores shown to be comparable with other non-clinical Western norms scores in females (Luce et al., 2008; Quick & Byrd-Bredbenner, 2013; Welch et al., 2011). Results suggest that UK male EDE-Q norms were marginally higher than previously published norms within non-clinical Western male samples (Lavender et al., 2010; Quick & Byrd-Bredbenner, 2013; Reas, Øverås, & Rø, 2012), yet future research should look to investigate whether such norms differ statistically between countries. Indeed, higher male norms in the present study support previous research which suggests that differences in EDE-Q norms may exist between countries within Western cultures (Mond et al., 2006), and further highlights the need to provide normative data within different countries, in order to

provide an empirical context to interpret an individual's score within a certain country. Greater investigation of cultural and international differences in ED psychopathology is particularly important to undertake amongst males, given the paucity of EDE-Q norm research in males compared with females (Lavender, Brown, & Murray, 2017).

In line with our hypothesis, a CFA failed to support the original four-factor structure (Fairburn & Beglin, 1994) for both sexes. This is consistent with previous literature in clinical and non-clinical samples (Aardoom et al., 2012; Forsén Mantilla et al., 2017; Rand-Giovannetti et al., 2017; White et al., 2014), which reinforces the argument that the original, theoretically derived EDE-Q structure may lack empirical support. Additionally, the current results replicated Allen et al. (2011) by failing to support alternative three, two, and one-factor models, which have been previously shown to provide the best fit to the data amongst clinical and non-clinical samples (Peterson et al., 2007; Penelo et al., 2013; Pennings & Wojciechowski, 2004). This would suggest that alternative structures which include all 22 attitudinal items are sub-optimal in capturing ED psychopathology, in both females and males. Indeed, Allen et al. (2011) concluded that a brief, 8-item, single-factor structure provided an adequate fit for the EDE-Q. Whilst this simplified structure may not capture the richness and complexity of ED psychopathology, our study conducted EFA of the full 22 attitudinal items in a large sample of female and males, to assess which items did not clearly load onto any one factor.

The EFA revealed that a three-factor model - *Shape and Weight Concern*, *Preoccupation and Eating Concern*, and *Restriction* - was the most appropriate fit to the data for both females and males, using an 18-item and 16-item model, respectively. Whilst such findings do not suggest the model should contain as few items as previous literature has proposed (Allen et al., 2011; Gideon et al., 2016; Grilo et al., 2015; Kliem et al., 2016), it is supportive in suggesting that the EDE-Q would benefit from a revised, briefer version than the

current 22-item attitudinal measure, whilst maintaining a comparable degree of assessment value. Indeed, newly-proposed EDE-Q scores are provided for all samples based on the respective three-factor structure for both sexes (see Supplementary Materials – Tables S8 & S9), which similarly reflects higher subscale and global scores amongst females compared with males. Given such differences between sexes, our results suggest that males may require a different, validated scoring of this measure to reflect the difference in ED psychopathology compared with females. Moreover, EFA results within the present study support previous research which combines Shape Concern and Weight Concern subscales (Rand-Giovannetti et al., 2017; Forsén Mantilla et al., 2017; Peterson et al., 2007), suggesting that individuals in the non-clinical population interpret shape and weight as closely associated within a generalized body concern. Indeed, Shape Concern and Weight Concern items accounted for approximately half of the model variance explained for both sexes. This is particularly important to consider with regard to the interpretation of normative EDE-Q data within the non-clinical population, given that scores on Shape Concern and Weight Concern are typically the highest of the original, four EDE-Q subscales (e.g. see Table 1). Consequently, a high proportion of individuals can display norm scores above the clinical cut-off (≥4; Mond et al., 2006) which presents a false positive in the number of cases falling within clinical significance. Indeed, within the present study, almost 1 in 4 non-clinical females scored above this clinical cut-off on Shape Concern alone (see Supplementary Materials Table S4). This suggests that clinical cut-off thresholds need to be reassessed, and further underpins the need to provide up-to-date normative data from both clinical and non-clinical populations, to statistically determine the accurate use of this measure as a clinical screening instrument (Mond et al., 2006).

Despite several studies undertaking EFA to determine a new EDE-Q factor structure (e.g. Forsén Mantilla et al., 2017; Machado et al., 2014; White et al., 2014; Darcy et al., 2013), to the authors' knowledge, only three studies have subsequently validated their own newly-

proposed EDE-Q structure within an independent sample (Friborg et al., 2013; Hrabosky et al., 2008; Kliem et al., 2016). In the present study, we aimed to assess our own newly-proposed factor structure across two independent, multi-site samples for both females and males. Given the congruency in EDE-Q norms with other Western samples, and validation of our newly-proposed three-factor structure in independent, multi-site samples of both student and non-student females, our results suggest that this factor structure could be applied to the general, non-clinical female population. Whilst similar validation was shown for non-student males, fit indices were marginally below the accepted CFA threshold for student males. This may be, in part, due to the smaller sample size in males ($N = 164$) influencing fit indices, however, it was ensured prior to analysis that the number of male participants across all samples provided sufficient statistical power to conduct a CFA (Bentler, 1995; Kline, 2014; Munro, 2005). Nevertheless, future research is encouraged to assess the newly-proposed factor structures for both females and males, which would further validate such models in wider non-clinical samples.

Further, high covariances were observed between error terms within the *Shape and Weight Concern* factor in both of the validated CFA models. This is likely to reflect the highly similar item wording within EDE-Q subscales, particularly within Shape Concern and Weight Concern subscales. At its most extreme, items within these subscales could be interpreted by respondents as qualitatively equivalent, with certain items only differing in the words "shape" or "weight" within an otherwise identical question (e.g. item 25 & 26). Indeed, this is supported by the EFA models in the present study which combine Shape Concern and Weight Concern subscales, in line with previous research in non-clinical samples (e.g. Rand-Giovannetti et al., 2017; White et al., 2014; Darcy et al., 2013). Such issues could lead to biases in factor loading estimates and subsequently affect the interpretability of EDE-Q subscale scores. Thus, a version of the EDE-Q which includes fewer items, particularly within Shape Concern and

Weight Concern subscales, without reducing the variance explained by the model is likely to provide a more reliable assessment measure.

Importantly, it is not suggested that the EDE-Q is without assessment and diagnostic value. However, the factor structure based on the 22 attitudinal items appears to cluster differently to the original, theoretically-proposed structure (Fairburn & Beglin, 1994). The factors extracted from the present EFA do closely resemble the originally proposed factors, but indicate that the EDE-Q may benefit from revision based on current body image ideals, using empirically driven methods. For example, items 6 *(Flat Stomach),* 10 (*Fear of Weight Gain*), and 12 (*Desire to Lose Weight*) were removed from both female and male EFA models. For males, this may reflect a more recent focus on leanness and muscularity in male body ideals (Pope, Phillips, & Olivardia, 2002), with less concern towards weight loss. In females, such poor item loading may be a consequence of the EDE-Q design, which may not fully capture the broader ED diagnoses under the new DSM-5 criteria (Vo, Accurso, Goldschmidt, & Le Grange, 2017). The EDE-Q was originally designed to assess anorexia nervosa (AN) and bulimia nervosa (BN) symptoms (Cooper, Cooper, & Fairburn, 1989). However, since its conception, ED diagnosis now incorporates new disorders characterised by different symptomatology. For example, binge eating disorder (BED) has only recently been introduced as a distinct diagnosis in the DSM-5, characterised by individuals eating large quantities of food and regularly experiencing loss of control over eating (American Psychiatric Association, 2013). Whilst select BED characteristics are captured in disordered eating behaviour items (e.g. objective binge episodes), the BED profile is not addressed in the attitudinal items which contribute to EDE-Q norm scores (Carrard, Lien Rebetez, Mobbs, & Van der Linden, 2015). This is particularly important when considering sex differences, given that, unlike AN and BN, BED affects males and females equally (Mitchison & Mond, 2015).

Nevertheless, whilst it is advised that methods of assessment in ED psychopathology should be revised in line with up-to-date diagnostic criteria, it is important to note that the above conclusions drawn in the present study are made based on non-clinical samples only. Previous research has provided a strong rationale for investigating such measures within the non-clinical population which delivers valuable information in promoting positive health-related outcomes (Mond et al., 2013), and provides reference data to map onto clinical ED assessment and diagnosis (Mond et al., 2009). However, future research should also investigate the factor structure and measurement invariance across clinical ED samples with differing symptomology, according to the latest DSM-5 criteria.

As noted in previous research (Mond et al., 2014; Reas et al., 2012), the sensitivity of the EDE-Q within male populations requires further evaluation. EDE-Q items focus largely on shape and weight concerns which are designed to assess a thinner body (e.g. *Restraint*). However, male body concerns are often based on reversed ideals, with a desire to increase muscularity and body mass (Lavender et al., 2017; Smith et al., 2017). Therefore, lower scores and ED prevalence amongst males may not be due to lower incidence of disordered eating attitudes *per se*, but rather that male body concerns are not adequately captured within EDE-Q items. Indeed, measurement invariance analysis of the previously proposed factor structures between sexes in the initial CFA (Phase 1) offers statistical support for females and males interpreting items in a conceptually different manner (see Supplementary Materials S5). These different conceptual interpretations may also be evident in disordered 'compensatory' behaviour responses, as 'excessive exercise' might reflect increased muscle mass for males, but increased weight loss for females (Lavender et al., 2010; Murray, Griffiths, & Mond, 2016). Thus, whilst male norms were lower than female norms in the present study, differences in the expression of ED symptomology may be better captured by accompanying assessment tools measuring body image concerns and ED psychopathology in males (Darcy et al., 2013;

Jennings & Phillips, 2017; Mond et al., 2014). Moreover, future studies must establish accurate, empirically derived clinical EDE-Q cut-offs (see Supplementary Materials S4), based on norms from clinical and non-clinical samples for females and males independently (Machado et al., 2014; Welch et al., 2011).

Studying ED vulnerability in student samples is of great importance, given the increased rates of body dissatisfaction (Berg, Frazier, & Sherr, 2009) and ED symptoms within this population (Lipson & Sonneville, 2017). Whilst the present study maintained a homogeneous sample in the initial CFA and subsequent EFA (Sample 1), the use of an exclusively student sample may have limited the generalizability of these findings to other non-clinical samples. However, we aimed to externally validate each of the newly-proposed factor structures in independent student (Sample 2) and non-student (Sample 3) samples, which is a key strength of our design. Moreover, the present pattern of findings is in line with previous research which investigates EDE-Q norms and factor structure in non-clinical samples across different countries (Quick & Byrd-Bredbenner., 2013; Allen et al., 2011; White et al., 2014; Forsén Mantilla et al., 2017).

However, limitations must be considered within the present study. Whilst efforts were made to improve the generalizability of our results across independent samples, our study employed convenience sampling which may not fully generalize to the general population. Thus, future research which assesses the newly-proposed factor in more heterogeneous samples is advised. Indeed, an aim within the present study was to maintain homogeneity in the age of our samples (see *Methods* section), to provide greater reliability when comparing EDE-Q norms and factor structures between females and males. However, whilst there is a large body of evidence investigating the age of onset for a clinical ED diagnosis (Micali, Hagberg, Petersen, & Treasure, 2013), and ED pathology over the lifespan (Cash & Smolak, 2011),

research which specifically investigates the impact of age towards EDE-Q scores is scarce (Rø, Reas, & Rosenvinge, 2012) and should be increasingly investigated to statistically explore the effects of age on ED pathology. Finally, each of the samples used in the present study had a significantly larger ratio of females compared with males. Whilst it was ensured that the number of participants across all samples provided sufficient statistical power to conduct a CFA and EFA (Kline, 1994; Bentler, 1995; Munro, 2005), it would be beneficial for future research to assess the newly-proposed three-factor model in a larger male sample.

In conclusion, the present study provides an important contribution to the wider EDE-Q literature as the first to provide EDE-Q norms of the original, four-factor structure for females and males within a non-clinical UK adult sample. Such data is valuable in providing an empirical context to appropriately interpret EDE-Q scores in non-clinical and clinical populations within the UK, and helps to facilitate comparisons in ED psychopathology with other countries and cultures. Consistent with previous research, EDE-Q norm scores were higher amongst females compared with males. However, psychometric assessment suggested that the original four-factor EDE-Q structure was sub-optimal for both sexes, with a three-factor structure shown to be the most appropriate fit to the data for both females and males, using an 18-item and 16-item model, respectively. Given the ongoing changes in ED diagnostic criteria since the introduction of the original EDE-Q measure, the psychometric properties should be reassessed in accordance with such developments, within both clinical and non-clinical samples. Moreover, given the established differences in body ideals between females and males, increased research is required in the non-clinical male population, to further determine whether EDE-Q items are interpreted qualitatively differently between sexes.

References

Aardoom, J. J., Dingemans, A. E., Slof Op't Landt, M. C. T., & Van Furth, E. F. (2012). Norms and discriminative validity of the Eating Disorder Examination Questionnaire (EDE-Q). *Eating Behaviors*, *13*(4), 305–309. https://doi.org/10.1016/j.eatbeh.2012.09.002

Allen, K. L., Byrne, S. M., Lampard, A., Watson, H., & 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

American Educational Research Association, American Psychological Association, & N. C. on M. in E. (1999). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical Significance Methods: A Comparison of Statistical Techniques. *Journal of Personality Assessment*, *82*(1), 37–41. http://doi.org/10.1207/s15327752jpa8201

Bazzini, D. G., Pepper, A., Swofford, R., & Cochran, K. (2015). How Healthy are Health Magazines? A Comparative Content Analysis of Cover Captions and Images of Women's and Men's Health Magazine. *Sex Roles*, *72*(5–6), 198–210. https://doi.org/10.1007/s11199-015-0456-2

Becker, A. E., Thomas, J. J., Bainivualiku, A., Richards, L., Navara, K., Roberts, A. L., … Striegel-Moore, R. H. (2010). Validity and reliability of a Fijian translation and adaptation of the eating disorder examination questionnaire. *International Journal of Eating Disorders*, *43*(2), 171–178. https://doi.org/10.1002/eat.20675

Bell, B. T., & Dittmar, H. (2011). Does Media Type Matter? The Role of Identification in Adolescent Girls' Media Consumption and the Impact of Different Thin-Ideal Media on

Body Image. *Sex Roles*, *65*(7), 478–490. https://doi.org/10.1007/s11199-011-9964-x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Bentler, P. M. (1995). EQS structural equations program manual. *Encino*, 83–100.

Berg, K. C., Frazier, P., & Sherr, L. (2009). Change in eating disorder attitudes and behavior in college women: Prevalence and predictors. *Eating Behaviors*, *10*(3), 137–142. https://doi.org/10.1016/j.eatbeh.2009.03.003

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

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, *154*, 136.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Sage.

Carrard, I., Lien Rebetez, M. M., Mobbs, O., & Van der Linden, M. (2015). Factor structure of a French version of the eating disorder examination-questionnaire among women with and without binge eating disorder symptoms. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *20*(1), 137–144. https://doi.org/10.1007/s40519-014-0148-x

Carter, J. C., Stewart, D. A., & Fairburn, C. G. (2001). Eating disorder examination questionnaire: Norms for young adolescent girls. *Behaviour Research and Therapy*, *39*(5), 625–632. https://doi.org/10.1016/S0005-7967(00)00033-4

Cash, T. F. (2004). Body image: Past, present, and future. *Body Image*, *1*(1), 1–5. https://doi.org/10.1016/S1740-1445(03)00011-1

Cash, T. F., Morrow, J. A., Hrabosky, J. I., & Perry, A. A. (2004). How has body image changed? A cross-sectional investigation of college women and men from 1983 to 2001. *Journal of Consulting and Clinical Psychology*, *72*(6), 1081–1089. https://doi.org/10.1037/0022-006X.72.6.1081

Cash, T. F., & Smolak, L. (2011). *Body image: A handbook of science, practice, and prevention*. Guilford Press.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of- Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902

Cooper, Z., Cooper, P. J., & Fairburn, C. G. (1989). The validity of the eating disorder examination and its subscales. *British Journal of Psychiatry*, *154*(JUN.), 807–812. https://doi.org/10.1192/bjp.154.6.807

Dakanalis, A., Zanetti, M. A., Riva, G., & Clerici, M. (2013). Psychosocial moderators of the relationship between body dissatisfaction and symptoms of eating disorders: A look at a sample of young Italian women. *Revue Europeene de Psychologie Appliquee*, *63*(5), 323–334. https://doi.org/10.1016/j.erap.2013.08.001

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

Fairburn, C. G., & Beglin, S. (1994). Assessment of eating disorders: interview or self- report questionnaire? *Int J Eat Disord*, *16*(4), 363–370. https://doi.org/10.1002/1098-108X(199412)16:4

Fairburn, C. G., & Cooper, Z. (1993). The Eating Disorder Examination (12th edition). In C. G. Fairburn & G. T. Wilson (Eds.). In *Binge eating: Nature, assessment, and treatment* (pp. 317–360). New York, NY, US: Guilford Press.

Forsén Mantilla, E., Birgegård, A., & Clinton, D. (2017). Factor analysis of the adolescent

version of the Eating Disorders Examination Questionnaire (EDE-Q): results from Swedish general population and clinical samples. *Journal of Eating Disorders*, *5*(1), 19. https://doi.org/10.1186/s40337-017-0140-8

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 (g) and primary factors. *International Journal of Methods in Psychiatric Research*, *22*, 195–203. https://doi.org/10.1002/mpr.1389

Gideon, N., Hawkes, N., Mond, J., Saunders, R., Tchanturia, K., & Serpell, L. (2016). Development and psychometric validation of the EDE-QS, a 12 item short form of the eating disorder examination questionnaire (EDE-Q). *PLoS ONE*, *11*(5), 1–19. https://doi.org/10.1371/journal.pone.0152744

Grilo, C. M., Crosby, R. D., Peterson, C. B., Masheb, R. M., White, M. A., Crow, S. J., … Mitchell, J. E. (2010). Factor structure of the eating disorder examination interview in patients with binge-eating disorder. *Obesity*, *18*(5), 977–981. https://doi.org/10.1038/oby.2009.321

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

Guest, T. (2000). Using the Eating Disorder Examination in the Assessment of Bulimia and Anorexia. *Social Work in Health Care*, *31*(4), 71–83. https://doi.org/10.1300/J010v31n04_05

Heene, M., Hilbert, S., Harald Freudenthaler, H., & Bühner, M. (2012). Sensitivity of SEM fit indexes with respect to violations of uncorrelated errors. *Structural Equation Modeling*, *19*(1), 36–50. https://doi.org/10.1080/10705511.2012.634710

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis.

Psychometrika, *30*(2), 179–185.

Hrabosky, J. I., White, M. A., Masheb, R. M., Rothschild, B. S., Burke-Martindale, C. H., & Grilo, C. M. (2008). Psychometric evaluation of the eating disorder examination-questionnaire for bariatric surgery candidates. *Obesity (Silver Spring, Md.)*, *16*(4), 763–769. https://doi.org/10.1038/oby.2008.3

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.

Jacobson, N. S. Y., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19.

Jennings, K. M., & Phillips, K. E. (2017). Eating Disorder Examination–Questionnaire (EDE–Q): Norms for a Clinical Sample of Males. *Archives of Psychiatric Nursing*, *31*(1), 73–76. https://doi.org/10.1016/j.apnu.2016.08.004

Kaiser, H. F. (1961). A note on Guttman's lower bound for the number of common factors 1. *British Journal of Statistical Psychology*, *14*(1), 1–2.

Karazsia, B. T., Murnen, S. K., & Tylka, T. L. (2017). Is Body Dissatisfaction Changing Across Time ? A Cross-Temporal Meta-Analysis. *Psychological Bulletin*, *143*(3), 293–320. https://doi.org/10.1037/bul0000081

Kelly, N. R., Cotter, E. W., Lydecker, J. A., & Mazzeo, S. E. (2017). Missing and discrepant data on the Eating Disorder Examination Questionnaire (EDE-Q): Quantity, quality, and implications. *Eating Behaviors*, *24*, 1–6. http://doi.org/10.1016/j.eatbeh.2016.11.002

Kliem, S., Mößle, T., Zenger, M., Strauß, B., Brähler, E., & Hilbert, A. (2016). The eating disorder examination-questionnaire 8: A brief measure of eating disorder psychopathology (EDE-Q8). *International Journal of Eating Disorders*, *49*(6), 613–616. https://doi.org/10.1002/eat.22487

Kline, P. (2014). *An easy guide to factor analysis*. Routledge.

Lavender, J., Brown, T. A., & Murray, S. B. (2017). Men, Muscles, and Eating Disorders: an Overview of Traditional and Muscularity-Oriented Disordered Eating. *Current Psychiatry Reports*, *19*(6), 1–7. https://doi.org/10.1007/s11920-017-0787-5

Lavender, J., De Young, K., & Anderson, D. (2010). Eating Disorder Examination Questionnaire (EDE-Q): Norms for undergraduate men. *Eating Behaviors*, *11*(2), 119–121. https://doi.org/10.1002/erv.2324

Lipson, S., & Sonneville, K. (2017). Eating disorder symptoms among undergraduate and graduate students at 12 U.S. colleges and universities. *Eating Behaviors*, *24*, 81–88. https://doi.org/10.1016/j.eatbeh.2016.12.003

Luce, K. H., Crowther, J. H., & Michele Pole, M. (2008). Eating Disorder Examination Questionnaire (EDE-Q): Norms for Undergraduate Women. *European Eating Disorders Review*, *22*(6), 439–442. https://doi.org/10.1002/erv.2324

Machado, P. P. P., Martins, C., Vaz, A. R., Conceição, E., Bastos, A. P., & Gonçalves, S. (2014). Eating disorder examination questionnaire: Psychometric properties and norms for the Portuguese population. *European Eating Disorders Review*, *22*(6), 448–453. https://doi.org/10.1002/erv.2318

Mantilla, E. F., & Birgegard, A. (2016). Eating disorder examination questionnaire: Norms and clinical reference data from adolescent boys and girls in Sweden. *Psychiatry Research*, *239*, 156–162. https://doi.org/10.1016/j.psychres.2016.03.022

Micali, N., Hagberg, K. W., Petersen, I., & Treasure, J. L. (2013). The incidence of eating disorders in the UK in 2000-2009: findings from the General Practice Research Database. *BMJ Open*, *3*(5), 3–10. https://doi.org/10.1136/bmjopen-2013-002646

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., Hall, A., Bentley, C., Harrison, C., Gratwick-Sarll, K., & Lewis, V. (2014). Eating-disordered behavior in adolescent boys: Eating disorder examination questionnaire norms. *International Journal of Eating Disorders*, *47*(4), 335–341. https://doi.org/10.1002/eat.22237

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). Eating Disorder Examination Questionnaire (EDE-Q): Norms for young adult women. *Behaviour Research and Therapy*, *44*(1), 53–62. https://doi.org/10.1016/j.brat.2004.12.003

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2009). Comparing the health burden of eating-disordered behavior and overweight in women. *Journal of Women's Health*, *18*(7), 1081–1089.

Mond, J. M., Mitchison, D., & Hay, P. (2013). Prevalence, impairment in quality of life, and implications for prevention and health promotion. *Current Findings on Males with Eating Disorders*, *195*.

Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., Rodgers, B., Morgan, J. F., … Mitchell, J. E. (2008). Screening for eating disorders in primary care: EDE-Q versus SCOFF. *Behaviour Research and Therapy*, *46*(5), 612–622. https://doi.org/10.1016/j.brat.2008.02.003

Munro, B. H. (2005). *Statistical methods for health care research* (Vol. 1). lippincott williams & wilkins.

Murray, S. B., Griffiths, S., & Mond, J. M. (2016). Evolving eating disorder psychopathology: Conceptualising muscularity-oriented disordered eating. *British Journal of Psychiatry*, *208*(5), 414–415. https://doi.org/10.1192/bjp.bp.115.168427

Parker, K., Mitchell, S., O'Brien, P., & Brennan, L. (2015). Psychometric evaluation of disordered eating measures in bariatric surgery patients. *Eating Behaviors*, *19*, 39–48. https://doi.org/10.1016/j.eatbeh.2015.05.007

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

Payne, L. O., Martz, D. M., Tompkins, K. B., Petroff, A. B., & Farrow, C. V. (2011). Gender Comparisons of Fat Talk in the United Kingdom and the United States. *Sex Roles*, *65*(7), 557–565. https://doi.org/10.1007/s11199-010-9881-4

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), 1–11. https://doi.org/10.1371/journal.pone.0083245

Pennings, C., & Wojciechowski, F. L. (2004). The Eating Disorder Questionnaire (EDE-Q): Netherlands normative scores for anorexic patients and a non-eating disorder control group. *Gedragstherapie*, *37*, 293–301.

Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., … Grange, D. le. (2007). Psychometric Properties of the Eating Disorder Examination-Questionnaire: Factor Structure and Internal Consistency. *The International Journal of Eating Disorders*, *40*(2), 386–389. https://doi.org/10.1002/eat

Pope, H., Phillips, K. A., & Olivardia, R. (2002). *The Adonis complex: How to identify, treat and prevent body obsession in men and boys*. Touchstone.

Quick, V. M., & Byrd-Bredbenner, C. (2013). Eating Disorders Examination Questionnaire (EDE-Q): Norms for US college students. *Eating and Weight Disorders*, *18*(1), 29–35. https://doi.org/10.1007/s40519-013-0015-1

Rand-Giovannetti, D., Cicero, D. C., Mond, J. M., & Latner, J. D. (2017). Psychometric Properties of the Eating Disorder Examination–Questionnaire (EDE-Q): A Confirmatory Factor Analysis and Assessment of Measurement Invariance by Sex. *Assessment*, 107319111773804. https://doi.org/10.1177/1073191117738046

Reas, D. L., Øverås, M., & Rø, Ø. (2012). Norms for the Eating Disorder Examination

Questionnaire (EDE-Q) Among High School and University Men. *Eating Disorders*, *20*(5), 437–443. https://doi.org/10.1080/10640266.2012.715523

Rø, Ø., Reas, D. L., & Rosenvinge, J. (2012). The impact of age and BMI on Eating Disorder Examination Questionnaire (EDE-Q) scores in a community sample. *Eating Behaviors*, *13*(2), 158–161. https://doi.org/10.1016/j.eatbeh.2011.12.001

Robinson, E., & Aveyard, P. (2017). Emaciated mannequins: a study of mannequin body size in high street fashion stores. *Journal of Eating Disorders*, *5*(1), 13. https://doi.org/10.1186/s40337-017-0142-6

Smith, K. E., Mason, T. B., Murray, S. B., Griffiths, S., Leonard, R. C., Wetterneck, C. T., … Lavender, J. M. (2017). Male clinical norms and sex differences on the Eating Disorder Inventory (EDI) and Eating Disorder Examination Questionnaire (EDE-Q). *International Journal of Eating Disorders*, *50*(7), 769–775. https://doi.org/10.1002/eat.22716

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). Using multivariate statistics.

Ullman, J. B. (2001). A first course in structural equation modeling.

Villarroel, A. M., Penelo, E., Portell, M., & Raich, R. M. (2011). Screening for eating disorders in undergraduate women: Norms and validity of the spanish version of the eating disorder examination questionnaire (EDE-Q). *Journal of Psychopathology and Behavioral Assessment*, *33*(1), 121–128. https://doi.org/10.1007/s10862-009-9177-6

Vo, M., Accurso, E. C., Goldschmidt, A. B., & Le Grange, D. (2017). The Impact of DSM-5 on Eating Disorder Diagnoses. *International Journal of Eating Disorders*, *50*(5), 578–581. https://doi.org/10.1002/eat.22628

Watkins, M. W. (2005). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*, *5*(2), 8.

Welch, E., Birgegard, A., Parling, T., & Ghaderi, A. (2011). Eating disorder examination questionnaire and clinical impairment assessment questionnaire: General population and

clinical norms for young adult women in Sweden. *Behaviour Research and Therapy*, *49*(2), 85–91. https://doi.org/10.1016/j.brat.2010.10.010

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

Table 1: *Descriptive Data - Means (Standard Deviations) for original, four-factor EDE-Q subscales and global score, for female (N=851) and male (N=224) students (Sample 1).*

	Total (*N*=1075)	Females (*N*=851)	Males (*N*=224)	*t*	*p*	Cohen's *d*
Age	19.89 (1.98)	19.77 (1.73)	20.34 (2.69)	3.011	.003	.252
BMI	22.69 (4.07)	22.60 (4.11)[a]	23.04 (3.88)[b]	1.411	.158	.110
Restraint	1.30 (1.33)	1.37 (1.34)	1.05 (1.25)	-3.218	.001	.246
Eating Concern	0.94 (1.08)	1.03 (1.11)	0.60 (0.84)	-6.342	<.001	.436
Shape Concern	2.34 (1.62)	2.51 (1.58)	1.69 (1.59)	-6.905	<.001	.518
Weight Concern	1.94 (1.57)	2.10 (1.57)	1.31 (1.39)	-7.391	<.001	.535
EDE-Q Global	1.63 (1.25)	1.75 (1.25)	1.16 (1.11)	-6.896	<.001	.500

Note: *p* values corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995).

BMI: Body Mass Index.

[a] Females (*N*= 816); [b] Males (*N*= 216).

Table 2: *Fit statistics for four models of EDE-Q data in female (N=851) and male (N=224) students (Sample 1).*

		No. of items	χ^2 (df)	*p*	χ^2 /df	RMSEA	GFI	AGFI	NFI	CFI
Four Factor	Females	21*	2301.732 (183)	<.001	12.578	.117	.771	.711	.823	.835
	Males	21*	762.575 (183)	<.001	4.167	.119	.746	.679	.773	.816
Three Factor	Females	22	2648.098 (206)	<.001	12.855	.118	.753	.696	.807	.819
	Males	22	872.926 (206)	<.001	4.238	.120	.731	.670	.754	.799
Two Factor	Females	22	2878.955 (208)	<.001	13.841	.123	.727	.668	.790	.802
	Males	22	943.298 (208)	<.001	4.535	.126	.704	.640	.735	.779
One Factor	Females	22	3557.472 (209)	<.001	17.021	.137	.667	.597	.741	.752
	Males	22	1062.709 (209)	<.001	5.085	.135	.664	.594	.701	.743

Note: χ^2: chi-square; df: degrees of freedom; RMSEA: Root Mean Square Errors of Approximation; GFI: Goodness of Fit Index; AGFI: Adjusted Goodness of Fit Index; NFI: Normed Fit Index; CFI: Comparative Fit Index.

* = item 8 removed from analysis of four-factor model due to Heywood cases, thus 21 items were entered for the CFA.

Table 3: *Pattern Matrix of PAF analysis of female* (*N*=851) *(18 items) and male* (*N*=224) *(16 items) EDE-Q items (Sample 1).*

	Females (*N* = 851)				**Males (*N* = 224)**		
	Shape and Weight Concerns	**Preoccupation and Eating Concern**	**Restriction**		**Shape and Weight Concerns**	**Preoccupation and Eating Concern**	**Restriction**
	Factor 1	Factor 2	Factor 3		Factor 1	Factor 2	Factor 3
(27) Discomfort seeing body	**.921**	-.017	-.056	*(26)* Dissatisfaction with shape	**.919**	-.100	.066
(26) Dissatisfaction with shape	**.902**	-.107	.065	*(28)* Avoidance of exposure	**.874**	.090	-.119
(28) Avoidance of exposure	**.888**	-.062	-.038	*(22)* Importance of weight	**.859**	-.084	-.002
(25) Dissatisfaction with weight	**.853**	-.066	.084	*(25)* Dissatisfaction with weight	**.807**	-.029	.117
(23) Importance of shape	**.746**	.117	-.025	*(23)* Importance of shape	**.803**	-.050	.115
(24) Reaction to prescribed weighing	**.732**	.045	-.120	*(27)* Discomfort seeing body	**.782**	.152	-.074
(22) Importance of weight	**.721**	.153	-.019	*(24)* Reaction to prescribed weighing	**.778**	.011	-.185
(11) Feelings of fatness	**.669**	-.011	.196	*(11)* Feelings of fatness	**.588**	.040	.253
(20) Guilt about eating	**.480**	.286	.094	*(19)* Eating in secret	-.134	**.723**	-.078
(7) Preoccupation with food, eating or calories	-.063	**.849**	-.126	*(20)* Guilt about eating	.143	**.625**	.132
(8) Preoccupation with shape or weight	.170	**.689**	-.066	*(7)* Preoccupation with food, eating or calories	-.077	**.611**	.058
(9) Fear of losing control over eating	.007	**.660**	.095	*(8)* Preoccupation with shape or weight	.162	**.567**	.063
(5) Empty stomach	-.056	**.609**	.185	*(21)* Social eating	.159	**.528**	-.086
(2) Avoidance of eating	-.067	**.463**	.202	*(1)* Restraint over eating	-.013	-.103	**.891**
(19) Eating in secret	.181	**.406**	-.039	*(3)* Food avoidance	-.085	.054	**.817**
(3) Food avoidance	-.001	-.034	**.884**	*(4)* Dietary Rules	.015	.051	**.629**
(1) Restraint over eating	.025	-.018	**.864**				
(4) Dietary Rules	-.023	.107	**.692**				
Eigenvalue	8.98	1.73	1.26		7.84	1.58	1.37
% of variance	49.67	9.62	6.98		48.97	9.84	8.57
Cronbach's Alpha	.94	.82	.87		.94	.78	.80

Note: Values in bold represent highest loadings which comprise the respective factor.
EDE-Q item number (Fairburn & Beglin, 1994) in italicized parentheses.

Table 4: *Fit statistics for newly-proposed three-factor models of EDE-Q data in student (Sample 2) and non-student (Sample 3) females and males.*

			No. of items	χ^2 (df)	*p*	χ^2 /df	RMSEA	GFI	AGFI	NFI	CFI
Student sample	**New Three Factor**	Females (*N*=489) (without modification)	18	970.22 (132)	<.001	7.35	.114	.833	.784	.866	.882
		Females (*N*=489) (with modification)	18	375.16 (122)	<.001	3.13	.066	.923	.890	.948	.964
	New Three Factor	Males (*N*=164) (without modification)	16	307.30 (101)	<.001	3.04	.112	.818	.756	.796	.851
		Males (*N*=164) (with modification)	16	222.08 (98)	<.001	2.27	.088	.864	.811	.853	.911
Non-student sample	**New Three Factor**	Females (*N*=561) (without modification)	18	1306.17 (132)	<.001	9.90	.126	.798	.738	.844	.857
		Females (*N*=561) (with modification)	18	445.02 (114)	<.001	3.90	.072	.924	.886	.947	.924
	New Three Factor	Males (*N*=170) (without modification)	16	325.10 (101)	<.001	3.22	.115	.811	.745	.803	.854
		Males (*N*=170) (with modification)	16	135.37 (94)	.003	1.44	.051	.901	.858	.918	.973

Note: χ^2: chi-square; df: degrees of freedom; RMSEA: Root Mean Square Errors of Approximation; GFI: Goodness of Fit Index; AGFI: Adjusted Goodness of Fit Index; NFI: Normed Fit Index; CFI: Comparative Fit Index