Data Screening and EFA Assignment

**Abstract:**

Research in past decades has underscored the importance of body dissatisfaction (BD) in the development and maintenance of disordered eating and health practices (e.g., restrictive eating, exercise habits). Given that 90% of individuals with eating disorders are female, research has focused primarily on female BD. However, male BD has increased sharply over the past decade, and has been linked to maladaptive health behaviors (e.g., steroid use, excessive exercise). Although the presentation of BD contrasts greatly between males and females, many current measures do not delineate between features of male and female BD and are not sensitive to core features of male BD (e.g., muscularity and bulk). Measures specified for males attempt to focus on these core features, but incorrectly assume that these features are equally important for all males. Therefore, this study provided an evaluation of four measures of male BD using item response theory to provide an alternative method of analysis and scoring for a more complete picture of male BD. 841 participants completed the study. By examining the response patterns in the data, the following research questions were answered: 1) What is the level of discriminability of each item and which are the most informative? 2) Are there item biases along the spectrum of BD? 3) Are these scales equivalent measures of BD and are equivalent short-form measures possible? Results also provided evidence for future research on developing a computer-adaptive test (CAT) for BD.

**Data:**

**Question group 1:** Participants rate these questions from 1 (extremely dissatisfied) to 5 (extremely satisfied) about different body parts.

1. Weight
2. Leanness of shoulders
3. Leanness of arms (e.g., biceps/triceps)
4. Muscularity of arms (e.g., biceps/triceps)
5. Leanness of stomach/abdomen
6. Muscularity of stomach/abdomen
7. Leanness of chest/upper torso
8. Muscularity of chest/upper torso
9. Leanness of back
10. Muscularity of back
11. Muscularity of neck
12. Overall body build
13. Overall leanness of body
14. Overall level of body's muscularity

**Question group 2:** Participants rate these questions from 1 (never) to 5 (always) about how they feel about their bodies.

1. I think my body should be leaner
2. I am concerned that my stomach is too flabby
3. I feel dissatisfied with my overall body build
4. I think I have too much fat on my body
5. I think my abs are not thin enough
6. Has eating sweets, cakes, or other high calorie food made you feel fat or weak?
7. Have you felt excessively large and rounded (i.e., fat)?
8. Have you felt ashamed of your body size or shape?
9. Has seeing your reflection (e.g., in a mirror or window) made you feel badly about your size or shape?
10. Have you been so worried about your body size or shape that you have been feeling that you ought to diet?
11. I think my body should be leaner
12. I think I have too much fat on my body
13. Eating sweets, cakes, or other high calorie foods makes me feel fat
14. I feel excessively fat
15. Seeing my reflection (e.g., in a mirror or window) makes me feel badly about my body fat

**Question group 3:** Participants rate these items from 1 (strongly agree) to 5 (strongly disagree) about how they feel about their bodies.

1. I am happy with how much muscle I have compared to how much fat I have
2. Other people think I have a good body
3. I am a good weight for my height
4. I am hesitant to take my shirt off in public because people will look at my body
5. I have thoughts of dissatisfaction towards my body
6. I think I have a generally attractive body
7. My body looks healthy
8. The shape of my body is one of my assets
9. My body is sexually appealing to others

**Data screening:** Include the relevant output for each question.

1. Accuracy:
   1. Check for typos and problems with the dataset.
      1. Show that the data has scores that are accurate or that are inaccurate.
      2. Fix any typos in the data by deleting only that point.
   2. Reverse code question 1 group so that high scores on these items indicate body dissatisfaction.
2. Missing data:
   1. Include a table that shows the missing data by person (table of missingness by participant).
   2. Exclude participants who are missing too much data from the rest of the analysis.
   3. Include the percent/proportion of data missing by variable.
   4. Run *mice* to replace the data when appropriate.
   5. Include a summary of the data to show the missing data has been replaced.
3. Outliers
   1. Create Mahalanobis distance scores.
   2. Include a summary of the Mahalanobis scores or a summary of the number of outliers (true/false table).
   3. Exclude all multivariate outliers.
4. Additivity/Multicollinearity
   1. Are any of the variables too correlated?
   2. Run a correlation table (do not need to include).
   3. Use the symnum function to look at the data to see if any of the correlations are too high.
5. Normality
   1. Include the skew and kurtosis values for your variables.
   2. Include one univariate histogram.
   3. Include the multivariate histogram.
   4. Is the data normal?
6. Linearity
   1. Include the QQ plot for linearity.
   2. Is the data linear?
7. Homogeneity and Homoscedasticity
   1. Include a residual scatterplot of the data.
   2. Is the data homogeneic?
   3. Is the data homoscedastic?

**EFA:** Include the relevant output for each question.

1. Before you start:
   1. Include output for both KMO and Barlett’s test.
   2. What does the Kaiser-Meyer-Olkin statistic tell you about sampling adequacy? Give number and interpretation.
      1. How many participants do you have? Do you seem to meet the 10-15 people per item suggestion?
   3. What does Barlett’s test tell you? Give number and interpretation.
2. Number of factors:
   1. Theory suggests two factors.
   2. How many does the Kaiser criterion suggest?
      1. Include the sum of the eigenvalues for both criteria.
   3. How many does the scree plot suggest?
      1. Include the scree plot.
   4. How many does the parallel analysis indicate?
      1. Include the parallel analysis.
   5. Hint: If you see a set of loadings where one factor’s loadings **all** double load with another factor, you are overfactoring – a sign you’ve run too many factors.
3. Simple structure:
   1. Use maximum likelihood as the fitting estimation and direct oblimin for the rotation.
   2. Rounds of Loading Interpretation:
      1. Include the loadings for each round of analysis.
      2. Describe which questions were considered *bad* and why you excluded them.
      3. Continue this process until you achieve simple structure.
4. Adequate solution:
   1. Include the fit indices (RMSEA + CI, RMSR, CFI, TLI).
   2. Describe the fit indices.
   3. Include the reliabilities for each factor.
   4. Are the reliabilities any good?
   5. Label the factors based on the questions.
5. Create subscale averages for each participant based on your EFA results. You should have three new columns that are the *average score* of the factors created above.
   1. Include a summary of the means and standard deviations for these subscales.
6. Write up:
   1. Short description of the scale.
   2. Include a summary of data screening.
   3. List the type of analysis – rotation, fitting estimation, program used.
   4. Values on why you had enough people and an adequate set of correlations.
   5. Number of factors suggested you choose and why (scree, eigenvalues, parallel analysis)
   6. What questions you eliminated – why did you eliminate them? Go through the rounds one at a time.
   7. Simple solution table of the last round of loadings. (APA style!)
   8. Interpretation of the factors.
   9. Description of adequacy of solution – fit indices, reliabilities, and average/SD scores for the factors.