Qeios

Peer Review

Review of: "Development and Preliminary Psychometric Evaluation of the Height Dissatisfaction Scale"

Miriam Olivola[1]

1. University of Pavia, Italy

Article: Development and Preliminary Psychometric Evaluation of the Height Dissatisfaction Scale

Overall rating: 3 out of 5 stars

The study addresses a worthwhile gap in body image measurement and offers a promising starting point. My rating reflects the contribution of the scale and its encouraging initial results, alongside unresolved questions about statistical reporting and the evidence supporting its proposed scope. I recommend substantial revision before the central psychometric claims can be accepted with confidence.

Comments to the authors

I appreciated the clear rationale for developing a measure that can accommodate dissatisfaction with being either shorter or taller than desired. The inclusion of 433 participants with HDS data provides a useful basis for exploratory work, and the strong internal consistency estimates (alpha = .92; omega = .93) are encouraging. Providing the full instrument and scoring instructions is also helpful. The title appropriately describes the evaluation as preliminary, and the discussion already acknowledges the need for independent confirmation, test-retest reliability, and measurement invariance. My main suggestions concern what can be concluded from the present analyses and how readers can reproduce them.

1 Establish what the bidirectional evidence supports

The distinction between dissatisfaction with shortness and dissatisfaction in either direction is the main contribution of the paper. However, correlating HDS scores with signed height discrepancy (ideal minus actual height) does not directly test dissatisfaction irrespective of direction. A negative discrepancy identifies people who wish to be shorter, whereas a positive discrepancy identifies those who wish to be

 doi.org/10.32388/12ID5Q

taller. A positive correlation with this signed variable therefore cannot establish that greater dissatisfaction accompanies a larger discrepancy in both groups.

Please report how many participants wished to be taller, shorter, or remain the same height, with descriptive HDS results for each group. An association with absolute discrepancy would more directly address discrepancy magnitude, while group-specific analyses would help assess whether the pattern holds in both directions. Where subgroup sizes are too small for reliable inference, descriptive results and a clear qualification would be sufficient. The abstract and conclusion should describe the HDS as designed to accommodate both directions, while reserving claims of demonstrated equivalence for future validation.

2 Explain item development and construct coverage

The theoretical sources of the item pool are described, but the steps between those sources and the initial 14 items need more detail. Please explain who generated and reviewed the items, whether experts or intended respondents assessed their relevance and clarity, and whether pilot testing or cognitive interviews were undertaken. If these steps were not conducted, please acknowledge this limitation. This would make the use of the scale-development guidance cited in reference 33 more transparent.

The retained items also raise useful questions about construct boundaries. Item 12 assesses how height influences others' perceptions or treatment without specifying a negative effect; agreement could reflect awareness of an advantage rather than dissatisfaction. Item 6 specifically assesses appearing shorter, whereas the item about appearing taller was removed. Its loading of .41 and uniqueness of .83 suggest that it warrants particular attention. Please explain the rationale for this asymmetry and for treating these experiences as equally informative indicators of a common score. Future cognitive testing could explore these issues without requiring immediate, unvalidated changes to the final instrument.

3 Make the factor analysis reproducible

The KMO result and strong loadings for several core items support further examination of a general factor. Nevertheless, the analysis needs fuller reporting. Please identify the correlation matrix used, describe the distributions of the ordinal items, and explain the choice of maximum likelihood estimation. Given the acknowledgement of non-normality elsewhere, a sensitivity analysis using an appropriate ordinal approach would help establish whether the one-factor interpretation is stable.

Please also specify the parallel-analysis procedure, including the number of simulations, comparison criterion, and whether it used component or common-factor eigenvalues. Figure 1 contains negative later

eigenvalues, which can arise from a reduced common-factor matrix; the matrix and procedure should therefore be identified explicitly. The text lists the subsequent eigenvalues as .36, .32, and .19, but their correspondence with the plotted values is not clear. Please reconcile the figure, numerical results, and explanation of the 51.87% variance estimate. Reporting the initial 14-item solution, including the loadings and uniqueness of the two excluded items, would make the item-reduction decisions easier to evaluate. Independent CFA remains a reasonable next step rather than a requirement to obtain a new sample for this preliminary paper.

4 Reconcile sample sizes and significance tests

There is an important inconsistency between the reported correlations, p values, and sample sizes. Table 3 gives effective sample sizes of 398 to 431, while the HDS correlations with actual height and height discrepancy are reported as -.25 ($p = .003$) and .22 ($p = .015$). For samples in that range, conventional two-sided tests would ordinarily give substantially smaller p values for correlations of these magnitudes. Please check the original output and state the exact sample size for each correlation. If the p values reflect a different sample, adjustment, or testing procedure, explain this explicitly and align the text and table.

The missing-data description also needs clarification: listwise deletion is stated, but the varying correlation sample sizes suggest that different complete-case subsets may have been used. Please distinguish exclusions for entirely missing HDS responses from partial item nonresponse and missing comparison measures, and specify how scores were calculated in each case. For the dependent-correlation comparisons, identify the implementation, correlation inputs, and common complete-case sample used for each test. Explain how inference was adapted to Spearman correlations; a paired bootstrap could provide a useful sensitivity check. Please audit all reported z statistics and corresponding p values as part of this reconciliation.

5 Interpret the comparative validity findings cautiously

The stronger HDS associations with eating disorder symptoms and self-esteem are interesting, but some of that advantage may reflect overlapping content. HDS item 11 explicitly concerns self-esteem, and items 7 and 8 concern changing body fat or muscle mass. These items may increase the association with the chosen comparison measures independently of how well the HDS assesses height dissatisfaction.

Please discuss this possibility and, if feasible, repeat the relevant correlations after omitting the overlapping items as a sensitivity analysis. Such an analysis should be presented as a robustness check, rather than as validation of a new shortened scale. Stronger cross-sectional correlations do not, by

themselves, demonstrate superior validity or incremental prediction. The wording about predicting psychological outcomes should therefore be replaced with wording about contemporaneous associations. Self-reported actual and ideal height should likewise be described as self-reported height indices rather than objective measurements.

6 Clarify the comparison measures and sample description

The EDE-QS description on page 7 differs from the instrument developed by Gideon et al. (reference 34). The published short form uses four response categories scored 0 to 3; the manuscript describes five categories scored 0 to 4 and presents the anchors as frequency-based throughout. Please verify the questionnaire actually administered and its scoring. If this is a reporting error, correct it. If the response format was modified, describe the modification and its implications for interpreting the validity findings.

Please also report the numbers recruited through student courses and community sources, the outcome of the attention checks, and whether any further exclusions occurred. Clarify what identifying as Australian represents, as nationality and cultural background are not interchangeable. Pooled associations with actual height may partly reflect differences between women and men, so stratified or adjusted analyses would be informative if the available data permit. These would complement the already acknowledged limitations of the predominantly young, female sample; comprehensive invariance testing can appropriately remain a future research goal.

Further reporting corrections

The supplement clearly identifies the reverse-scored item and states that clinical cut-offs have not been established. Please add the explicit transformation for item 2 (7 minus the response), the possible mean-score range of 1 to 6, and a rule for incomplete questionnaires. Clarify whether Table 1 reports item 2 before or after reverse scoring, and provide descriptive statistics for the total HDS score. Confidence intervals for reliability coefficients and key correlations would help readers judge precision.

Please replace the OSF placeholder on page 15 with a working repository link and ensure that the statements about available data and materials accurately describe what readers can access. Finally, check reference 22 against the discussion of women's experiences on page 3: the reference listed is a qualitative study of men of Chinese ancestry, so its support for that particular statement needs clarification or correction.

Overall assessment

I see value in developing a dedicated measure of height dissatisfaction, especially one intended to include people who wish to be shorter. The present findings justify continued development, but the proposed breadth of the instrument and the comparative validity claims need more careful support. The priority is to resolve the numerical inconsistencies, document the analyses, and align the conclusions with the evidence available. With those revisions, the paper could make a useful contribution as an initial psychometric study.

References checked for this review

Boateng GO et al. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in Public Health, 6, 149. https://doi.org/10.3389/fpubh.2018.00149

Gideon N et al. (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), e0152744. https://doi.org/10.1371/journal.pone.0152744

Declarations

Potential competing interests: No potential competing interests to declare.