Reviewer's report

Title: A revised short version of the Compassionate Love Scale for Humanity (CLS-H-SF): Evidence from Item Response Theory Analyses and Validity Testing

Version: 1 Date: 05 Nov 2019

Reviewer: Tianlan Wei

Reviewer's report:

Thank you for this opportunity for me to review the manuscript titled "A revised short version of the Compassionate Love Scale for Humanity (CLS-H-SF): Evidence from Item Response Theory Analyses and Validity Testing." I have now carefully reviewed the article and will provide my opinions below.

Being an advocate of the application of IRT techniques in affective/personality tests myself, I applaud the authors' efforts in getting the first IRT-based analysis done for CLS-H. However, I have some major concerns regarding the IRT-based selection criteria in shortening the scale as well as the practical significance of the study. The selection criteria, in particular, would caution me against recommending this article for publication. I will detail my concerns below.

1. First of all, the primary purpose of this study, shortening the 21-item, psychometrically sound CLS-H, needs stronger justifications in the article. The scale already has a reasonable number of items and shouldn't be a burden to most of its potential respondents. Is it of any major importance to shorten it? The authors do discuss the weaknesses and limitations of Santa Clara brief compassion scale (SCBCS) along the way. But the main question should not be "what's wrong with SCBCS," but indeed "what's wrong with CLS-H?" Moreover, wouldn't a shortened version impair the "satisfactory content validity" (p. 4, L. 80) of the scale? In addition, I kind of see how giving a CLS-H "in healthcare, workplace, educational settings" (line 100) could take place, but I am not fully convinced. I would have liked to see a stronger justification for the usefulness of the scale. Administer it in educational/healthcare/workplace settings to what end? What would the employers do with this information???

2. Similarly, there's no theoretical justification for investigating the Differential Item Functioning across gender groups. While it seems intuitive that males and females may, on average, demonstrate differing levels of compassionate love, what makes it plausible that they interact with the CLS-H items differently?

3. A typo on p. 8, L. 202: it should be "NNFI and CFI &gt; .95," not "&lt;="

4. Methodological concerns:

4.1 To the best of my knowledge, IRT is barely used for the purpose of shortening a scale (though Rasch or other IRT models are often used in scale development). This is
because the IRT parameters are estimated with the entire data set as input data; therefore, removing any one item would result in changes in the item parameters (even though IRT is "sample-independent"). Thus, the "better than average information' criterion" (p. 9. L. 213) is post hoc, arbitrary, and METHODOLOGICALLY VERY CONCERNING. I also read the cited source for this criterion (Huang et al., 2017), only to find that Huang et al. also did it arbitrarily, without citing any credible source to endorse this important methodological decision.

4.2 More importantly, item information function (IIF) is mostly proportional to the discrimination (a) parameter but less related to the threshold (b) parameters. As we may see from Table 1, the items with above average information are also items with the highest discrimination (a) values, while some of the selected items do not seem to cover a reasonably wide range of the latent trait being measured (e.g., b6 for Item 3 is barely 1.0, meaning that this item is incapable of capturing the differences among individuals high in compassionate love).

4.3 For DIF across gender, more detailed information should be given in regard to the selection of anchor items in the several rounds of DIF analysis. Also, should the alpha level be adjusted/corrected to control familywise Type I error rate? Even Item 12 (p = .008) perhaps shouldn't be flagged as demonstrating significant DIF if we consider this. More importantly, what is the theoretical reasons behind the DIF across gender groups? Such should be discussed to provide implications and guidance for future compassion love research. This is only briefly mentioned in the Discussion section with unclear language (i.e., "Additionally, investigating gender differential item functioning . . . that appeared to be biased against gender . . ." should be phrased as "biased against 'males.'")

4.4 For reliability estimates, the information presented in the article seems a bit confusing. The authors mention Cronbach's alpha (p. 13, L. 288) but is it the CTT-based alpha or a different reliability estimate? Cronbach's alpha is total, raw score based, and I cannot see how it is estimated at different theta levels (or maybe it's a method I don't know). Also, it seems meaningless to compare the alphas of the original and shortened versions because the alpha coefficient also acts as a function of the item number k. I would like more information about the IRT/theta-based reliability estimates.

4.5 Figure 1 should be re-created to reflect the loss of information of the shortened version (vertical axes should be aligned; or consider combining the TIFs in one graph).

5. Some minor writing/stylistic issues:

5.1 I suggest a verification that the abbreviated journal names are what's required. They impede the ability to locate the article (which is the main purpose of the references section after all).

5.2 One stylistic comment: the authors use the expression "in order to" very many times throughout the article. I suggest revising all instances into simply "to."
5.3 Some sentences exhibited linguistic quirks/typos (for example, line 90: "this short scale was obtained adopting poor item selection criteria" should probably read: "this short scale was created BY adopting…"). I suggest another round of edits for clarity.

5.4 One sentence was really hard to follow and requires revision (line 65: "These responses from the individual allow one to pay attention to the needs and miseries of others, thus facilitating an individual to support others around them intimately.") I am honestly not sure what this means.

5.5 Lastly, a minor point, but stood out prominently to me: make sure to have consistent spacing (one space before and one after) whenever "=" is used throughout the paper, especially in the Measures section.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.