Author’s response to reviews

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

Authors:
FRANCESCA CHIESI (francesca.chiesi@unifi.it)
Chloe Lau (clau263@uwo.ca)
Donald Saklofske (don.saklofske@uwo.ca)

Version: 3 Date: 10 Feb 2020

Author’s response to reviews:

Dear Dr. Clayton,

We were pleased that our manuscript "A revised short version of the Compassionate Love Scale for Humanity (CLS-H-SF): Evidence from Item Response Theory Analyses and Validity Testing" (PSYO-D-19-00146R1) is being considered for publication in BMC Psychology, subject to a minor revisions and response to the comments raised by one of the reviewers.

We enclosed a copy of the manuscript revised following the BMC editorial Policies as well as our point-by-point responses to the additional comments made by Reviewer 1.

We would like to take this opportunity to thank you and the reviewers for helping us to improve the quality of our manuscript.

Kind regards,

Corresponding Author of the Manuscript

Following the Editorial Policies,

1. We described the role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript (p.23, line 508-511).

2. We uploaded the figures as separate, individual files in the file inventory and we removed the titles and legends now reported at p. 21, after the References section.

3. We provided a final, clean version without tracked changes.
Reviewer #1:

Comment 1. One of the justifications given is regarding the content validity of the scale: the authors argue that the content validity has been a concern questioned by theorists, but their confirmatory factory analysis (CFA) yielded results that strongly endorse the unidimensionality of the scale. What are the authors' thoughts on this?

Response: Thank you for this comment that gives us the opportunity to better explain our point. Content validity is about the extent to which the items on a test are fairly representative of the entire domain the test seeks to measure. Content validity is most often measured by relying on the knowledge of people who are familiar with the construct being measured. These subject-matter experts are usually asked to evaluate if the concept that is being measured is adequately covered by the item content. Thus, statistical analyses are not performed, but just a qualitative assessment made by experts. Factor analysis is usually used to assess factorial (internal, structural) validity that examines the extent to which the underlying structure of a scale reflects the construct, which might be defined by a single latent dimension or multiple dependent/independent dimensions. Thus, the concerns regarding the original Compassionate Love Scale was about item content and not regarding its factorial structure. For this reason, excluding some items might help in improving the scale.

Comment 2. Edelen & Reeve's (2007) IRT-based item selection procedure (can be found on p. 14 of the article) may still seem a little "post hoc" and exploratory, without any clear-cut guidelines (this is not a problem per se). I wonder if there's any more recent publications (e.g., 2015-present) shedding lights on this issue.

Response: To the best of our knowledge, this is a reference work about using IRT in shortening a scale, especially because a very didactical and practical approach was adopted to illustrate this IRT application. Additionally, shortening a scale is a step-by-step procedure made of many decisions that are driven by the test specific characteristics and the goal we’d like to achieve. Thus, we agree that this procedure is partially exploratory and, as a consequence, in literature we find different approaches and there are not clear-cut or general guidelines that fit every situation.

Comment 3. When the two graphs in Figure 2 were recreated to be align with each other using the same scale, we could still observe some loss of the test information, as well as increased standard error. To make the following statement convincing, the authors do need to cite credible resources to show that the advantages of the shortened scale may outweigh the loss of information and increased stand errors, preferable in a more object, quantitative way.

Response: Thank you to the reviewer for describing the issue regarding loss of information and increased standard errors. Given that each item information function would add to the overall test information function to increase the amount of information and decrease standard error, we would expect that there is a loss information and increased standard error for all short forms compared to a long form. Thus, our goal is to minimize the amount of information loss rather than preventing loss of information altogether. Again, to the best of our knowledge there are not
credible source to cite that give us clear cut-offs to decide if the loss of information (reliability) is acceptable or not. However, we deem that Table 2 provides a practical and quantitative way to give objective data about that. Moreover, we partially modified the sentence (see p.12, line 278-280) using additional quantitative information.