LETTER OF REVIEWERS

Reviewer D:
Recommendation: Accept Submission

Relevance: High
Novelty: High
Presentation and writing: High

Comments for authors:
It is suggested that after the discussion, you incorporate the limitations of the study as well as the implications for future research, as well as the implications in other areas such as clinical, educational and social.

Reviewer E:
Recommendation: Revisions Required

Relevance: High
Novelty: Very high
Presentation and writing: Moderated

Comments for authors:
Abstract:
1. The abstract should be structured according to the guidelines of the journal: Introduction, Aim, Method, Results, and Conclusion.

Introduction:
2. The introduction seems to me to be very well done and clear. I have no comments on this part.

Methods:
3. Participants: It is unclear to me if the authors have used a sufficient number of participants to perform their analysis. While I consider the sample to be large, it is not clear that they have sufficient statistical power to conduct their analysis based on the complexity of their model, the number of items, and the proposed CFI cut-off. I suggest adding a statistical power calculation for the instrument. For example, you can analyze this calculator: https://wnarifin.shinyapps.io/ss_sem_cfi_unequal/
For example: “For unidimensional dimensions with five items, a CFI of 0.95 is expected, with mean factor loadings of 0.5, a significance level of 0.05, and a power of 80%, so a sample size of 300 participants is expected for these analyses.”
This is a model that can be adapted by the authors based on the results of the calculator.

4. Instruments: It was a bit confusing to read the instruments part because the authors talked about the instrument itself and then about the construction process. I would suggest talking about the construction process in the pilot test first: “For the construction of the instrument, a process by phases was carried out beforehand. In the first phase, a bank of items for each dimension was written, based on the theoretical characteristics of each one of them. They were tested using an evaluation procedure by expert judges, with a total of 6 participating. Afterward, a pre-test was carried out on 63 people to put the items to the test and see if the participants had any difficulty answering them, as well as to determine the functioning of the items, remaining in total 5 items per dimension, except in the power dimension, with 4. Once confirmed, the final sample was obtained to carry out the analyses.”
And then talk about the instrument as such.

5. In the instrument section, it was unclear if there are correlated and independent dimensions; or if the instrument has a higher order multidimensional structure such as second order, bifactor, ESEM bifactor. I suggest adding this information to make subsequent interpretations of the results clearer.

6. In the data analysis section, it was not clear to me which method was used to select the number of dimensions, the Kaiser criterion or parallel analysis. I suggest the latter, as the former has been criticized for being subjective.

7. The EFA should indicate the estimate used (ULS, ML, other). Please add.

8. I ask the authors to reconsider whether the ML estimate is the best estimator for the nature of their variables (categorical-ordinal). Evidence suggests that the WLSMV estimate performs better for ordinal items (Likert 1-5): https://www.tandfonline.com/doi/pdf/10.1080/10705511.2014.937669
I suggest rerunning the CFA analysis using this estimator. This estimator may yield better goodness-of-fit indices.

9. I suggest adding the SRMR or WRMR to the CFA.

10. It is unclear how the authors performed this analysis, please specify: “A Confirmatory Factor Analysis was also carried out with all the scales, both for the answers corresponding to the couple and to those corresponding to one of the other people, to see if they worked as a one-dimensional test to obtain a global score.”

11. In the results section, the authors state that they evaluated a total dimension, but in the data analysis subsection, there is no information on how the multidimensional models were analyzed. I suggest adding the process of scoring the multidimensional models. For example, how was it evaluated that there is no overlap between dimensions? For example, with latent correlations between dimensions <0.80. It is also necessary to point out that no adjustment for correlated errors was made.

12. The study is very interesting. I suggest adding an analysis of invariance between groups so that the instrument can also be used to make comparisons between groups, for example, between men and women. This would help a lot with the internal validity of the instrument.

13. I would suggest adding an assessment of a congeneric internal consistency coefficient that does not assume tau equivalence, such as the omega coefficient. Since the study did not assess the tau-equivalence of the model and only presented alpha, this may introduce bias in the reliability estimate.

14. An "Ethical aspects" subsection should be added, stating that the protocol was approved by an ethics committee and the approval code. It should also state that all participants completed an informed consent form and what ethical care was required.

15. It is unclear how participants were included in the EFA and CFA. The suggestion is always to select a random sample of participants for the EFA (half of the participants) and the remaining randomly selected participants for the CFA (the other half of the sample). I would recommend reporting this, as it is not suggested to explore and confirm in the same sample, as this would introduce bias.

16. Within the Data Analysis section, it is suggested to add which data analysis packages were used. For example, R Studio was used with the Lavaan, SemPlot, etc. packages. If R or Mplus was used, it is suggested to add the script as supplementary material to make it easier to replicate the results.

Resultados:
17. However, the EFA report is somewhat confusing in parts. In tables 3 and 5, the KMO index indicates that there is sample adequacy for the EFA. This means that there is a sufficient number of participants. In both tables, the rotated factor loadings of the dimensions are not shown. Although this may be a bit lengthy, it helps to better understand how the items are organized into the 8 proposed dimensions. As each factor has its own KMO, I understand that an individual analysis has been carried out for each

factor. If this is the case, I would suggest that this be more clearly stated in the section and analysis and briefly repeated in the results section. Also, note that all EFAs for each factor were unidimensional.
On the other hand, the authors say that the multidimensional model fits 8 dimensions, and an analysis is presented in both tables with KMO values. For multidimensional models, I would recommend presenting supplementary material with the factor loadings of the complex model.

18. The authors point out: “The following values of the indexes used have been taken as a reference of the goodness of fit of the models: RMSEA≤.10, TLI>.95 and CFI≥.95 (Schreiber et al., 2006) and the quotient χ2/gl < 5 (Abad et al., 2011).” This should be in the Methods section. Please move it.

19. I would suggest that the SRMR estimator could be added to Tables 6 and 7.

20. I think that Tables 6 and 7 can be merged as they both assess the same variables but in different groups.

21. Figures 1 to 7 could be combined into a single figure to make it easier to see. It would help a lot to make it easier for the reader. It would also be good to have a figure of the multidimensional model with dimensions 1 to 7, where you can see if they are correlated or if there is a general factor (bifactor, second order, other).

22. I would suggest adding a measurement invariance analysis to make the manuscript more robust.

Discussion:
23. It is suggested to add a subsection on limitations and strengths. Also, add a sub-section on clinical implications to clarify the applicability of the findings.

Minor comments:
24. I think the clarity of the text could be improved as this statement is unclear:
“The selection criteria was that the person surveyed had a partner at the moment of the survey, selecting volunteers to participate in it.”

RESPONSE LETTER

Dear reviewer and editor,

The team of authors appreciates the review conducted and the recommendations provided by the reviewer, as they represent an improvement in the quality of the article. Below, we outline the response to the various comments, hoping they address them adequately. We have resubmitted the article with the changes marked in blue.

Abstract:
1. The abstract should be structured according to the guidelines of the journal: Introduction, Aim, Method, Results, and Conclusion.

***Answer**: We have adapted the abstract following these recommendation, and we added it in Spanish language.*

Introduction:
2. The introduction seems to me to be very well done and clear. I have no comments on this part.

***Answer:** We appreciate your comment.*

Methods:
3. Participants: It is unclear to me if the authors have used a sufficient number of participants to perform their analysis. While I consider the sample to be large, it is not clear that they have sufficient statistical power to conduct their analysis based on the complexity of their model, the number of items, and the proposed CFI cut-off. I suggest adding a statistical power calculation for the instrument. For example, you can analyze this calculator: https://wnarifin.shinyapps.io/ss_sem_cfi_unequal/
For example: "For unidimensional dimensions with five items, a CFI of 0.95 is expected, with mean factor loadings of 0.5, a significance level of 0.05, and a power of 80%, so a sample size of 300 participants is expected for these analyses."
This is a model that can be adapted by the authors based on the results of the calculator.

***Answer:** We believe this tool is very interesting, but the article is extensive, so we have added a more concise version at the end of the sample section (page 10, paragraph 2).*

4. Instruments: It was a bit confusing to read the instruments part because the authors talked about the instrument itself and then about the construction process. I would suggest talking about the construction process in the pilot test first: "For the construction of the instrument, a process by phases was carried out beforehand. In the first phase, a bank of items for each dimension was written, based on the theoretical characteristics of each one of them. They were tested using an evaluation procedure by expert judges, with a total of 6 participating. Afterward, a pre-test was carried out on 63 people to put the items to the test and see if the participants had any difficulty answering them, as well as to determine the functioning of the items, remaining in total 5 items per dimension, except in the power dimension, with 4. Once confirmed, the final sample was obtained to carry out the analyses."
And then talk about the instrument as such.

***Answer:** Thank you for the recommendation. We have included additional information about the test construction and the development of a pilot phase on page 10, paragraph 1.*
We have also moved it to the beginning of the Method section to provide better clarification before delving into the sample section (page 9, paragraph 2).

5. In the instrument section, it was unclear if there are correlated and independent dimensions; or if the instrument has a higher order multidimensional structure such as second order, bifactor, ESEM bifactor. I suggest adding this information to make subsequent interpretations of the results clearer.

***Answer:**A change has been made to improve the analyses and results, and a second-order factor has been included to obtain a total score for the scale. To do this, a second-order confirmatory*

factor analysis (CFA) with diagonally weighted least squares (DWLS) extraction method was tested. It is found that the second-order fit is very similar to that of independent factors. This has been added to the manuscript in the analysis and results sections (page 12, paragraphs 3 and 4; page 17, paragraph 1).

6. In the data analysis section, it was not clear to me which method was used to select the number of dimensions, the Kaiser criterion or parallel analysis. I suggest the latter, as the former has been criticized for being subjective.

Answer:*The Kaiser criterion K1 and the scree plot have been used, although extraction of 7 dimensions was forced to assess its adequacy to the theoretical model (page 12, paragraph 3).*

7. The EFA should indicate the estimate used (ULS, ML, other). Please add.

Answer:*We have used the Principal Axis Factoring method for parameter estimation, which is suitable when the variables do not meet normality. This has been added to the analysis description (page 12, paragraph 3).*

8. I ask the authors to reconsider whether the ML estimate is the best estimator for the nature of their variables (categorical-ordinal). Evidence suggests that the WLSMV estimate performs better for ordinal items (Likert 1-5): https://www.tandfonline.com/doi/pdf/10.1080/10705511.2014.937669
I suggest rerunning the CFA analysis using this estimator. This estimator may yield better goodness-of-fit indices.

Answer:*It has been considered to use the DWLS extraction method, suitable for a sample of over 200 subjects, and the analysis has been repeated.*

9. I suggest adding the SRMR or WRMR to the CFA.

Answer:*The SRMR index has been calculated and added to the manuscript (table 6).*

10. It is unclear how the authors performed this analysis, please specify: “A Confirmatory Factor Analysis was also carried out with all the scales, both for the answers corresponding to the couple and to those corresponding to one of the other people, to see if they worked as a one-dimensional test to obtain a global score.”

Answer:*We have clarified this point (page 12, paragraph 4, page 17 paragraph 1 and table 6). Two CFA analyses are tested using the DWLS extraction method, observing the goodness-of-fit indicators CFA, TLI, SRMR, and RMSEA in all cases. The Chi-Square value has been disregarded to test the model, as with high subject volumes, the Chi-Square indicator often yields statistically significant results. Given the ineffectiveness of this criterion for analyzing fit, the aforementioned goodness-of-fit indices are alternatively used.*

Second-order models are also tested, both in the CFA conducted on the couple and in different individuals within the couple. In all cases, the model fits are adequate, and consideration is given to using second-order factor models for their explanatory capacity and utility for testing. This has been incorporated into the manuscript.

11. In the results section, the authors state that they evaluated a total dimension, but in the data analysis subsection, there is no information on how the multidimensional models were analyzed. I suggest adding the process of scoring the multidimensional models. For example, how was it evaluated that there is no overlap between dimensions? For example, with latent correlations between dimensions <0.80. It is also necessary to point out that no adjustment for correlated errors was made.

Answer:*This aspect has been incorporated by including a model with both first and second-order factors.*

12. The study is very interesting. I suggest adding an analysis of invariance between groups so that the instrument can also be used to make comparisons between groups, for example, between men and women. This would help a lot with the internal validity of the instrument.

Answer:*This contribution is very enriching for the study. Multi-group invariance has been tested for men and women, after excluding cases of individuals who identified as non-binary (n = 6).*

Invariance models are tested with the second-order CFA model, both in couple scores and scores of other members. In all cases, metric, scalar, and strict invariance are met, assuming that the scale can be used invariantly between men and women. The criterion for analyzing invariance is the change in goodness-of-fit indices (CFI, TLI, RMSEA, SRMR), with Δ < .01. This has been added to the article (page 13, paragraph 1, page 18, table 7).

13. I would suggest adding an assessment of a congeneric internal consistency coefficient that does not assume tau equivalence, such as the omega coefficient. Since the study did not assess the tau-equivalence of the model and only presented alpha, this may introduce bias in the reliability estimate.

Answer:*The internal consistency tests for total scores and subscales are reanalyzed, both in couple members and other members. In all cases, as suggested by the reviewer, McDonald's omega is examined. The results are satisfactory (Omega > .7), with no significant changes compared to the alpha values, which are also provide (page 13, paragraph 1, and page 17, tables 6 and 7).*

14. An "Ethical aspects" subsection should be added, stating that the protocol was approved by an ethics committee and the approval code. It should also state that all participants completed an informed consent form and what ethical care was required.

Answer:*This information has been added to the article (Page 12, paragraph 2).*

15. It is unclear how participants were included in the EFA and CFA. The suggestion is always to select a random sample of participants for the EFA (half of the participants) and the remaining randomly selected participants for the CFA (the other half of the sample). I would recommend reporting this, as it is not suggested to explore and confirm in the same sample, as this would introduce bias.

Answer:*A random split-sample approach was employed, dividing the sample into two halves. Exploratory factor analysis (EFA) was conducted with the first half, and confirmatory factor analysis (CFA) was performed with the second half. This has been added to the manuscript (page 12, paragraph 3).*

16. Within the Data Analysis section, it is suggested to add which data analysis packages were used. For example, R Studio was used with the Lavaan, SemPlot, etc. packages. If R or Mplus was used, it is suggested to add the script as supplementary material to make it easier to replicate the results.

Answer: *We added this information: the software used to carry out the analyzes was IBM SPSS version 25 for the EFA and JASP version 0.18.3 for the CFA (page 13, last sentence of paragraph 1).*

Results:

17. However, the EFA report is somewhat confusing in parts. In tables 3 and 5, the KMO index indicates that there is sample adequacy for the EFA. This means that there is a sufficient number of participants. In both tables, the rotated factor loadings of the dimensions are not shown. Although this may be a bit lengthy, it helps to better understand how the items are organized into the 8 proposed dimensions.
As each factor has its own KMO, I understand that an individual analysis has been carried out for each factor. If this is the case, I would suggest that this be more clearly stated in the section and analysis and briefly repeated in the results section. Also, note that all EFAs for each factor were unidimensional.
On the other hand, the authors say that the multidimensional model fits 8 dimensions, and an analysis is presented in both tables with KMO values. For multidimensional models, I would recommend presenting supplementary material with the factor loadings of the complex model.

Answer: *The rotated factor loadings have been included in Table 2. The Kaiser-Meyer-Olkin (KMO) index has been included only for the multidimensional model. This has been clarified in the article (page 12, paragraph 3).*

18. The authors point out: “The following values of the indexes used have been taken as a reference of the goodness of fit of the models: RMSEA≤.10, TLI>.95 and CFI≥.95 (Schreiber et al., 2006) and the quotient $\chi 2/gl < 5$ (Abad et al., 2011).” This should be in the Methods section. Please move it.

Answer: *Thank you, we have moved it to the corresponding section (page 12, paragraph 2).*

19. I would suggest that the SRMR estimator could be added to Tables 6 and 7.

Answer: *Thank you, we have included it.*

20. I think that Tables 6 and 7 can be merged as they both assess the same variables but in different groups.

Answer: *We fully agree with this recommendation. We have merged tables 6 and 7 into table 6 (page 17), with the former table 7 now corresponding to the invariance measure.*

21. Figures 1 to 7 could be combined into a single figure to make it easier to see. It would help a lot to make it easier for the reader. It would also be good to have a figure of the multidimensional model with dimensions 1 to 7, where you can see if they are correlated or if there is a general factor (bifactor, second order, other).

Answer: *Indeed, the graphical model could be significantly reduced. By incorporating the second-order model, a clearer model has emerged. Additionally, due to its large dimensions, we have decided to include only the diagram and add an additional table (figure 1 and table 8) with parameter estimation values for the main scale (couple analysis) responded regarding the couple and regarding other individuals. A new version of the Classification of the Couple scale was created following this format (Figure 2 and table 9). The result is a total of two figures (pages 19 and 20).*

22. I would suggest adding a measurement invariance analysis to make the manuscript more robust.

Answer: *Done.*

Discussion:

23. It is suggested to add a subsection on limitations and strengths. Also, add a sub-section on clinical implications to clarify the applicability of the findings.

Answer: *We have included this section at the end of the discussion (page 24, paragraphs 2 and 3).*

Minor comments:

24. I think the clarity of the text could be improved as this statement is unclear:
"The selection criteria was that the person surveyed had a partner at the moment of the survey, selecting volunteers to participate in it."

Answer: *We added this information in sample section (page 9, paragraph 1).*

We have also included improvements in the article, such as making clearer the distinction between the 7 dimensions that make up the main scale, which we have called couple analysis, and the additional scale that maintains its name as Classification of the couple. This change can be observed on page 10, paragraph 2, and in the different tables. We also added the corresponding new references (page 23 and 27).