Reviewer's report

Title: Searching for the optimal number of response alternatives for the distress scale of the Four Dimensional Symptom Questionnaire

Version: 2 Date: 27 Aug 2018

Reviewer: Adrian Alacreu-Crespo

Reviewer's report:

This manuscript evaluates two ways of coding the scale of distress of 4DSQ using the IRT and provide convergent, divergent and predictive validity to this two ways of code the scale. In my opinion the paper is novel, well written and provide information useful for the clinical practice. In this sense, I think that paper would be publish with a slight review. Following, I explain some issues that would be revise before publication:

1. First of all, please provide all the information of 4DSQ in instruments, also, make clear that you are going to analyze only the distress scale (provide statistics from past research as alpha’s), although appear in introduction please explain the two ways of coding, and explain what of the four measures of 4DSQ from MISS sample were used to evaluate predictive validity. It was baseline measure? One of the most important concerning things for me is, why you not evaluate the precision from the other scales? This would improve paper a lot, and the other scales are more or less compatible to test convergent, divergent and predictive validity, although of course it would be more interesting if there exist results with the HML or the BDI for depression scale or the STAI for Anxiety scale.

2. I don’t understand the rationale of use the Neuroticism scale of NEO-FFI as discriminant validity. Divergent validity implies that the construct evaluate (distress) not are relate with a construct that a priori should’nt be related with. In this regard, distress and neuroticism was relate in the past (eg. Ploubidis & Frangou, 2011) Moreover your results show relationships between this constructs. Then, that would be another measure of convergent validity?

3. In the BPL, why you discard the physical functioning statements? Can you provide the rationale in the instruments?

4. With respect to UBOS, Table 3 not show the short-term results. UBOS were asked only in the last follow up? Please explain it in the instruments section.
5. In page 12 line 10 you said Table A1 and Table A2 to refer the item score frequencies from first (0-2 coding) and second (0-4) coding; but you commit a mistake, is Table A2 and Table A3. Regarding this tables, it is one of the most important results in the paper, I think that this tables would appear in the results section and not in appendix.

6. Regarding convergent-validity section (page 13) please provide p values. Furthermore, take into account the coment 2 for interpret this section.

7. Regarding predictive validity, I have some concerns and proposals. First of all, is the first time in paper when short-term/long-term labels appear, and it people would be lost when they read it. I conclude that short-term prediction speaks about the six-month follow up and long-term about the twelve-month follow up. Then, I think that this issue would be fixed describing more this differentiation in the participants description. As I said in coment 1, what measure of 4DSQ was utilized for perform predictive validity? That would appear before. Another issue, did you compared 4DSQ with the two-month follow up measures? Moreover, as you explained in participants section this participants have elevate levels of distress, then this participants are distressed from the begining and there are not participants without stress. In this sense it would be useful first if you provide the descriptive information from the baseline and the follow ups of the different test. Moreover, it would be interesting calculate the changes in distress and the other scales (follow up - Baseline) and see if distress decline corresponds with decline in the other scales.

8. Regarding the interpretation of results, in the abstract and the conclusion in page 17 you conclude that recoding should be avoided because decrease the precision of test. However, your results not provide complete evidence for say it. In one hand, the model fit for the three items alternative is better than the five items. But, on the other hand, the "sometimes" (that is the 1 in the 3 items model) category is redundant, and three items show less predictive validity in a couple of scales compared to 5-items model. In this regard, that seems mixing results to me, and seems that 3-items is also useful. I´m agree with the recommendation, but the coment in the abstract is to heavy for me, and I propose authors say that is a recomendation instead of should be that is more imperative.

9. In the discussion you speak about the convergent, discriminant and predictive validity, but you not explain anything about the direction of relationships and the interest of this relationships in the live of patients, it would be interesting interpret this relationships in the discussion.
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.

Yes

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:

Acceptable

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