Reviewer’s report

Title: Explaining Differential Item Functioning focusing on the crucial role of external information – An example from the measurement of adolescent mental health

Version: 1 Date: 22 Jul 2019

Reviewer: James Greenwood-Lee

Reviewer's report:

The major concern I outlined was not adequately addressed, in that there were neither major revisions to the text to address the comment, nor a direct response to the comment. The latter would be warranted if the if the comment wasn't a valid criticism due to my misunderstanding, in which case a response would potentially reveal changes that could be made to prevent similar misunderstandings from readers like myself who are not necessarily familiar with all of the technical aspects. The comment I refer to is:

"However, this last quote brings my main criticism forward, which is the lack of discussion around two critical issues. The first is causality. A hypothesis is put forward that post-menarche females are more likely to report stomach aches as a result of menstrual pain. The subsequent analyses support this hypothesis, however, the results do not establish causality; rather the results simply show that the data support the hypothesis. This brings me to the second point which is alternative hypotheses. A key assumption of the authors hypothesis is that "girls from the two groups (pre and post menarche) experience the same overall load of psychosomatic problems". Indeed a key step in establishing DIF is to condition on group membership where it is established that the groups are invariant in the latent construct of interest. However, what if this is not true. The social pressure on grade 9, 7 and 5 females are likely to be very different, and an alternative hypothesis is that grade 9 girls are subject to the greatest amounts of anxiety, which triggers them to respond in the survey as experiencing a higher rate of stomach aches. Now, Table 1 shows that menarche increases dramatically between grade 5, 7 and 9, which means menarche will subsequently be positively correlated with the reporting of stomach aches, but under this alternative hypothesis the appearance of DIF due to menarche is created by an ordinary group difference that we are interested in capturing. How do we reconcile these two alternative hypotheses? In particular, how is it that we're justified in assuming that "girls from the two groups experience the same overall load of psychosomatic problems" in the first place? Could one not argue that by making this assumption, we are assuming that there are no ordinary group differences, and therefore we are assuming that any difference observed between the groups must be created by our measurement tools? If this is the case, doesn't the whole thing crumble?".
Here I turn to a concern expressed by reviewer 2 which I think nicely summarizes my concern more generally: "One major concern I have is that the paper is very difficult to understand and doesn't provide enough guidance for application, which I think should be the main contribution of the manuscript. As BMC medical research methodology is about general methodology in medical research rather than specific for measure development, it is important for this manuscript to explain the issue clearly in plain language and provide a guidance to researchers who may be interested in understanding the issue and/or using the approaches in their own studies and these researchers are not necessarily familiar with all the methodology-specific terminology." I have to agree and believe that it should be rewritten to be more accessible.

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

Yes

**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.

Unable to assess

**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 recommend additional statistical review

**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