Reviewer's report

Title: Interprofessional assessment of medical students' competences with an instrument suitable for physicians and nurses

Version: 0 Date: 20 Nov 2018

Reviewer: Wolf Hautz

Reviewer's report:

Dear authors,

thank you for the opportunity to review your manuscript on Interprofessional assessment of medical students' competences.

I applaud you for conducting such an extensive assessment with students and assessors from three medical schools, which I believe to pose a rather substantial logistical challenge. What adds to the relevance of the paper is the fact that it is still very much unknown how competences should be assessed, especially those you refer to as "specific non-medical".

To further improve your paper, I have several suggestions.

Introduction:

Line 104ff: Please be more specific here. I would suggest to name a few of the conditions which need to be met to achieve satisfactory psychometric characteristics (including reliability) in an OSCE. I think the claim that "the students are assessed at different stations by more than one rater, mostly physicians, [...] leads to good reliability" per se is too broad. The AMEE guide you cite gives a number of conditions which should be met to achieve satisfactory reliability in an OSCE.

Lines 110ff: I would suggest differentiating between assessment purposes. While OSCEs are often used in summative exams, where reliability is important, I think there is a general agreement in the literature that work place based assessments should only be used for formative purposes (Frank Med Teach 2010, Schuwirth Med Teach 2011, Lörwald Med Teach 2017), precisely because they are rather unreliable.

Research question, line 133ff: Your introduction leads to the first aim of your study, namely the question if different assessor groups assess candidates differently. The second part (how confident assessors are in their ratings) however comes as a bit of a surprise. I think it's a quite innovative aspect to look at, but I would suggest to provide a little bit of a rational as to why you are interested in assessor confidence. I am not aware of any research in that field, but the literature on confidence judgements in learning and clinical reasoning (see eg. Cue utilization framework by Anique de Bruin, Med Educ 2017, or a recent review by Ilgen et al in ASHE 2018, the work of Asher Korriat or Eva & Regehr on self Monitoring) could provide starting
points. A rational as to why you assess confidence, together with a hypothesis, could also help to frame the discussion of this particular aspect.

One could for example hypothesize that when confident, raters tend towards bolder judgments (more towards the borders of the scale) whereas when unconfident (but forced to provide a score anyway) they might be more lenient (i.e. scoring towards the scales mean).

Methods:

Lines 153ff: I would suggest to explain how you came up with the 7 competencies you assess, in particular because in the introduction (line 97), you claim that "different competences should be assessed separately, to give raters the opportunity to focus on each competence individually".

I would further suggest extending the analyses. It is interesting that nurses score more lenient than the other assessors. One important question however is whether this simply reflects a shift of scales, or whether the rank order of students changes depending on the assessor. In other words: can we convert the supervisors score into a nurses score simply by multiplying it with a fixed factor and those that score high in the supervisors assessment also score high with the nurses, or are the best/worst performers actually different depending on assessor.

Results:

Line 191: Please substitute "p<0.05" with "all p<0.05".

Line 189: I was surprised to see that you calculated alpha for your assessment. If I understand correctly, you propose that your assessment measures 7 different competencies. Then why would you calculate only one alpha? In theory, the 7 competencies you assess should be (at least) partially be independent of each other (because otherwise, they would be just one competence). A factor analysis could tell.

Discussion:

As far as I understand, assessors are obliged to provide a score (unless they did not observe the respective behavior at all). It would be interesting (from my point of view), whether it is their confidence in their assessment that predicts their score or whether their score informs their confidence. Please see my comment on confidence above: I would suggest to extend the discussion of the relationship between score and confidence in it to the cues that may inform either.

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.
Yes

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

Yes

**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.

none

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