Reviewer’s report

Title: Validity and Reliability of the Spanish version of the Organizational Readiness for Knowledge Translation questionnaire (OR4KT)

Version: 0 Date: 18 Jun 2017

Reviewer: Rohit Ramaswamy

Reviewer's report:

Thank you for the opportunity to review this article. This paper continues to contribute to the important field of measuring organizational readiness for change and builds on the substantial work that has already been done in the careful development of the OR4KT instrument. The details on the conceptual underpinnings this instrument and the review of theories, frameworks and models that led to the initial development of concepts, dimensions and sub-dimensions have been published in this journal and elsewhere. So too have the Delphi studies that led to the English version of the instrument with 6 dimensions and associated 24 sub-dimensions. I have not read the companion paper that describes the process of translation and cultural adaptation of the English version to Spanish and French but am familiar with the process followed by the authors and with the care taken to ensure rigorous translation and adaptation. The careful attention to methodological detail is evident in this paper as well.

This paper is ambitious in the fact that it attempts to assess not just the reliability and the internal validity of the OR4KT Spanish version, but also concurrent and discriminant validity. The methodology for assessing the reliability and internal validity is sound, except for the use of the word "reproducibility". If I understand correctly, the ICC was calculated asking the same physicians to take the survey test "under similar circumstances". I believe this is a measure of "repeatability", not "reproducibility".

The description of the EFA methodology is clear as is the discussion of the results. However, it feels to me that trying to present complex results from the factor analysis and the results from the concurrent and discriminant validity analyses in the same paper has resulted in incomplete discussion about the implications of the factor analysis. There are many interesting findings from the analysis that have been left unexplored. For example, in the organizational climate dimension, the authors report that the last two sub-dimensions (communication about change and openness to change) formed a single sub-dimension. But a review of the items from the supplementary material in the appendix reveals that they explore different aspects of organizational climate. Items 7 and 8 ask about well functioning communication channels and freedom of staff to ask questions, while items 9 and 10 ask whether the clinic director is open to new practices and change ideas. Therefore the fact that the EFA indicated that these items loaded on a single factor would have benefited from an exploration of the item ratings to determine why this was the case. The results of the Leadership/participation dimension are would also have
benefited from a detailed exploration of item responses. The importance of leaders and champions and the existence of a strategic planning process have consistently emerged as criteria for implementation success (including in the recent review of PSV III using CFIR constructs by Martinez in BMC Family Practice) and if, as the authors state, the objective of the OR4KT is to use the instrument as a screening tool to identify organizations that will be successful in implementing complex chronic care interventions, then a more detailed analysis of the underlying factor structure than is currently presented in the paper will certainly help to better understand the psychometric properties of the instrument.

Indeed, the authors appear to get distracted from presenting a detailed description of the internal validity analysis by their eagerness to report the concurrent and discriminant validity results. The discriminant validity analysis is the weakest part of the paper. In Figure 1, only 4 out of the 19 practices (I thought there were 20 practices, but Figure 1 only has 19) have more than 50% of practitioners who have signed up for PSV III. This is too small a sample to assess any predictive capability of OR4KT, since so few of the practices appear to be ready in the first place. The ROC analysis seems to be a complex way of saying that: (i) the scores of all facilities ranged from 52.25 to 68.94, (ii) that none of the facilities where fewer than 50% of the physicians signed up had a score > 64.48 and (iii) three of these facilities had a higher score than one facility where > 50% of the physicians signed up. From this data, stating that 64.48 is the optimal cut-off score for discrimination is not credible. The fact that 33% of the facilities where fewer than 50% staff sign off (3 out of 9) have a higher score than an "engaged" facility, and that the difference between the cut-off score and the next lower score (which was achieved by a facility deemed "disengaged") is so small (e.g. 63.45 vs 64.48) that any conclusions about discriminant validity seem a stretch.

But this gets to a more fundamental question of what exactly the concurrent and discriminant validity analyses are trying to accomplish, and to the broader objectives of OR4KT itself. The premise behind the development of the instrument, as stated in the introduction to the paper, is to assess organizational readiness for "the implementation of evidence-informed health promotion interventions in primary care". But the current wording of the items cannot directly measure this because the refer more broadly to generic interventions and changes in the clinic. However, implementing a complex intervention such as PSV III is not the same as implementing an IT system in an organization where there is a standard and preferred mode of engagement of organizational personnel with the system. But in a program such as PSV III, the success is determined by the complex activities of a large number of actors (doctors, nurses, community members) and one of the key implementation strategies for PSV III is to allow local adaptation and participatory design (Sanchez, BMC Health Services Research, 2009). An instrument that had the capability to assess readiness to implement the particular interventions of PSV III in each setting has to have some way of introducing a description of these interventions into the items themselves. The current wording in the instructions "think about a change, intervention, new practice etc. in the context of an innovation to promote health living habits" is too general.
However, radical transformation of the items to suit each PSV III implementation could affect the psychometrics. Therefore a discussion about what aspects of readiness to implement a program such as PSV III can be captured by the current version of OR4KT and what research might be needed to explore customized adaptations of the items would be very valuable in this paper.

All of this is to say that it is important to have greater discussion about the appropriate analyses to assess concurrent, predictive or discriminant validity. The authors are correct in saying that we need to go beyond reliability and internal validity, but how to do so cannot be a haphazard process. In this paper, the assessment of correlation of OR4KT with SOAPC makes sense. The two instruments measure similar constructs. Since SOAPC was developed for primary care settings, provides confidence for the ability for OR4KT to be used in these settings as well. I do not know the theoretical underpinnings of the construction of SOAPC but OR4KT comes from a robust theoretical approach and therefore could be a stronger instrument to assess to organizational characteristics of primary care settings as they ready for change, and perhaps to assess changes in these characteristics after an implementation has taken place. However, as with the case of the discriminant validity analysis, the value of performing a concurrent validity analysis with the PPP doesn't make conceptual sense. This instrument measures the extent to which individual providers engage in health promotion behaviors, and this is different from organizational characteristics of primary health care settings. Concurrent validity reflects the extent to which a measure correlates with another validated instrument measuring the same construct, and it is not clear to me that PPP measures the same constructs as OR4KT. It feels to me that more research about adaptations of the items to more closely reflect PPP items, is needed if PPP is to used as a validating instrument for OR4KT.

In summary, this paper continues to strengthen the already strong analytics around OR4KT. But it is uneven in its concurrent and discriminant validity parts and incomplete in its analysis of internal validity. I suggest revisions that incorporate the following:

- a more detailed analysis and description of the factor analysis results, with an item analysis to explain variations from the theory, an a summary of the final theoretical structure of the Spanish version. Brief comparisons with the English (and French, if available) versions would be useful.

- A practical interpretation of the concurrent validity analysis with SOAPC with implications for the use of OR4KT as a substitute for SOAPC in assessing the organizational readiness of primary care settings to implement change
- Less emphasis or elimination of the concurrent validity analysis using PPP and the discriminant validity analysis, but a discussion on additional research that might be undertaken on the extent to which context specific adaptations of OR4KT might be needed to use it as an instrument to predict the capability of organizations to implement complex interventions such as PSVIII.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**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 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