Author’s response to reviews

Title: Developing a tool to measure satisfaction among health professionals in Africa

Authors:

Adama Faye (adama.faye@umontreal.ca)
Pierre Fournier (pierre.fournier@umontreal.ca)
Idrissa Diop (idyhygea@yahoo.fr)
Aline Philibert (aline.philibert@crchum.qc.ca)
Florence Morestin (Florence.Morestin@inpq.qc.ca)
Alexandre Dumont (alexandre.dumont@ird.fr)

Version: 3 Date: 5 May 2013

Author’s response to reviews: see over
To reviewers 2 and 3
As a second major compulsory revision was requested, we took this opportunity to review the strategy for the explanatory analyses. During our previous work we put more emphasis on the validation steps.

As a result the tool has better psychometric properties (CFI=0.9415; RMSEA=0.0508 with 90% CI: 0.0448 – 0.0569). Its structure is slightly modified: the tool has now 8 dimensions instead of 9 (similar than in the previous one). Explained variance (EFA) is now 72.8% and was 55.3% previously.

Reviewer 2
Major Compulsory Revisions - The author must respond to these before a decision on publication can be reached.

1. The authors have improved the quality of the manuscript since the first review and seem to have addressed most of the original comments. However, the methodology could benefit from additional revisions.

2. The original comment read as following: PCA used for making “adjustments to the final instrument” is described under the section on construct validity. PCA is not used for construct validation in general. If it is used for screening an initial set of items or extracting factors, then it should be described under the appropriate section. Moreover, the reason for conducting PCA as opposed to other methods of extracting factors (for example, parallel analysis) is somewhat unclear, and it would seem that authors do not provide sufficient justification.

   In response, the authors removed the sentence on PCA and state the following:
   "By performing EFA, we were able to identify the underlying factor structure.... From the EFA analysis we found that, of the 44 items, 43 were on their original dimension, with loadings coefficients ranging from 0.42 to 0.84."

   We provide a description of the extraction rules and the parameters used for EFA.

   These sentences was added in the method section:
   *The extraction method used was principal component analysis based on a correlation matrix. The model was optimized through an orthogonal rotation (Varimax with Kaiser normalization). The extraction rules we used for our 44 initials items were the following ones: an extraction value equal or higher to 0.5, a loading factor on a first factor of 0.5 or higher and on a second factor equal or lower than 0.35.*

The application of rules for factors and items selection is
How did they decide on the number of factors following the EFA? Did the authors look at eigenvalues? Screeplot? Was the only criterion in selecting items following EFA was the size of factor loadings? What about fit indices? Residual variance? It seems to me that it is still unclear to the reader the methods and reasoning for extracting the 9 factors, which is an important part of the development of the scale.

Presented as follows:
Following the EFA, 8 factors were retained: all of them had an eigenvalue greater than 1 (decreasing from 6.255 to 1.113) with a scree plot showing a discernible elbow from the 8th factor (9th factor eigenvalue = .685). Among the 44 initial items, 24 of them met the inclusion criterions (see loading coefficients in Table 2 and the component matrix in Appendix V). Percentage of total variance explained by the 8 factors was 72.8% (Table 2).

3. The authors have not provided the estimators used in EFA and CFA and these are considered to be important basic parameters of the model. The response to referred to fit indices.

EFA (explanatory factor analysis): in the previous version, we already provided the factor loadings of each item considered and the % of total variance explained by each factor. In the revised version we added the factor and item extraction approaches and confirmed the appropriateness of the PCA by assessing the measure of sampling adequacy (KMO-Kaiser-Meyer-Olkin) and the Bartlett sphericity test. We also optimized our model with the orthogonal rotation method (Varimax)
- KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy was performed. KMO was higher than .5 (.875)
- Bartlett sphericity test was statistically significant (.000)

We added this sentence (see also answer to comment 2):
A Kaiser-Meyer-Olkin measure of sampling adequacy higher than .5 (.875) and a statistically significant Bartlett sphericity test (.000) insured the appropriateness of the analysis.

CFA (confirmatory factor analysis): We tested one oblique model as the factors were correlated. The estimation procedure we used was maximum likelihood including input variance-covariance matrix. Latent variables variance was fixed at 1. Parameters were estimated with a robust standard error option (Satorra-Bentler).

We added these sentences:
The model resulting from EFA was tested as an oblique model given since the 8 factors were correlated between each other. The estimation procedure used was the maximum likelihood based on an input variance-covariance matrix. The fit indices we
used were the comparative fit index (CFI) [40] and the root mean square error of approximation (RMSEA) [41]. Robust estimates of standard errors were also evaluated (Satorra-Bentler).

We also added the following information about the statistical software we used:


4. It will be good to provide a short table with basic item distributions as it is commonly done in literature (means, std etc). It is also helpful in seeing if there is a particular ceiling effect. Done, given to the table size we provided an Appendix named Appendix IV. In this new table means, standard deviations and medians are provided for all the 44 items for the 2 sub samples and for the total sample.

**Minor Essential Revisions**

1. It would be useful if the authors cited a methods paper in the section on criterion validity to support their approach, specifically the following: "The score was divided into quintiles, with the top quintile considered as 'very satisfied'. Chi-squared testing was used for comparisons." Basically, it appears they generated factor scores, made them into binary variables and then used them as observables.

In most validation studies, the criterion validity assessment is based on the correlation with existing scales. We explained why we used another approach and cited a paper that used the same approach than us

The decision to use the top quintile as the most satisfied is our own choice. As the variables of the tested hypotheses are categorical we could have compared the means scores of the different categories. The two approaches give the same results.

We added this sentence and one reference: **Validity criterion is usually based on comparison between an existing scale and the one under development. In our case as such appropriate scales do not exist for our dimensions, for each of them we selected 2 criterions based on hypothesis suggested by an expert’s panel. Tso and al. used the same approach for criterion related validation of their patient satisfaction scale, they correlated the patient satisfaction scores with hypothetically related criterions (intended future reutilization and recommendation to others).**

Reviewer 3

The authors have responded satisfactorily to most of my comments but I have still one major and from my perspective compulsory comment. It is related to the participants originating in the QUARITÉ project. It is said that the participants (N=962) were working in 46 hospitals. Did they comprise the total staff or were they recruited from a greater population consisting of all employees and in that case what was the size of the total population and can the recruitment process from any perspective be considered having been selective? The importance of this question is increased by the very high participation rate which is so high (97.4%) that it is hard to believe that it can be reached in real life. On the other hand I am aware of the fact that the representativeness of the data is not in main focus in the developmental work of a survey scale but nevertheless, these basic facts should according to my opinion be clarified before potential publishing.

We thought that we had already answered to this previous comment by adding these sentences in our previous revision:

In the Study population and sample section:

- The final stage, conducted in Senegal and Mali involved the health professionals working in the 46 hospitals selected for the QUARITÉ trial conducted between 2007 and 2011 [35].
- For the purposes of this survey, the sample consisted of all physicians, nurses and midwives practising in the 46 health facilities in the QUARITÉ trial (962 persons).

In the Strengths and weaknesses section:

- While generalization is improved due to the sample’s diversity, the fact that the respondents all worked in maternity services could limit generalization to all health sectors. Given the mobility of health workers—with the exception of midwives—from one sector to another one, this limitation might not be considered a major constraint.

In order to provide additional information about the sample and the working places, we added these sentences in the Study population and sample section that are directly extracted from the paper [35]:

- The trial was conducted in 46 out of a total of 49 eligible referral hospitals—23 in Mali and 26 in Senegal—spread across both countries. A hospital was eligible for the trial if it had functional operating rooms and carried out more than 800 deliveries annually. Three eligible hospitals were excluded: two already had a structured program for carrying out maternal death audits before the project began, and for one other hospital, written consent was not provided by local authorities. The 46 included hospitals are representative of the existing health system in Senegal and Mali, taking into account the variety of the contexts (urban versus rural) and of the levels of care (primary versus secondary referral health facilities).
- For the purposes of this survey, the sample consisted of all physicians, nurses and midwives practising in the in patients maternity wards of the 46 health facilities in the QUARITÉ trial (962 persons).